

```
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 38
Db 133 FNNFTVSFWLRVPKVSASHLE 154

RESULT 12
US-10-295-074-59
; Sequence 59, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: 3 hTNF joined by glycine linkers and P2 and P30 introduced
US-10-295-074-59

Query Match 42.7%; Score 117; DB 14; Length 514;
Best Local Similarity 95.7%; Pred. No. 4.4e-06;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPKVSASHLE 37
Db 492 ALFNNFTVSFWLRVPKVSASHLE 514

RESULT 13
US-10-297-942-2
; Sequence 2, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-2

Query Match 42.3%; Score 116; DB 14; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSFWLRVPKVSASHLE 37
Db 64 LFNFTVSFWLRVPKVSASHLE 85

RESULT 14
US-10-295-074-47
; Sequence 47, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
```

```
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
; APPLICANT: Pharmexa A/S
; NAME/KEY: MUTAGEN
; LOCATION: (110)..(130)
; OTHER INFORMATION: Tetanus toxoid P30 epitope
; FEATURE:
; NAME/KEY: MUTAGEN
; LOCATION: (131)..(145)
; OTHER INFORMATION: Tetanus toxoid P2 epitope
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(109)
; OTHER INFORMATION: hTNF amino acids 1-108
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (146)..(194)
; OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-47

Query Match 42.0%; Score 115; DB 14; Length 194;
Best Local Similarity 91.7%; Pred. No. 2.6e-06;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 PSLFNNFTVSFWLRVPKVSASHLE 37
Db 107 PEGFNNFTVSFWLRVPKVSASHLE 130

RESULT 15
US-10-297-942-10
; Sequence 10, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-10

Query Match 41.6%; Score 114; DB 14; Length 158;
Best Local Similarity 70.6%; Pred. No. 2.8e-06;
Matches 24; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 17 FNNFTVSFWLRVPKVSASHLE 42
Db 41 FNNFTVSFWLRVPKVSASHLE 74

Search completed: March 10, 2004, 10:25:49
Job time : 39.2996 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 14.6304 Seconds
(without alignments)

309.015 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSGPSLKLKLS.....HRLGVGSLHWSYGLRXP 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	29.0	282	P00376	cell fusion glycop
2	72	29.0	282	P00388	cell fusion glycop
3	72	29.0	534	JU0274	cell fusion glycop
4	72	29.0	546	S47300	gene F protein - r
5	72	29.0	550	E48556	cell fusion glycop
6	72	29.0	553	VGNZMV	cell fusion glycop
7	71	28.6	546	VGNZRK	cell fusion glycop
8	71	28.6	546	S47305	gene F protein - r
9	66.5	26.8	552	S47034	cell fusion protei
10	66	26.6	542	JQ2223	cell fusion glycop
11	66	26.6	546	VGNZRL	cell fusion glycop
12	66	26.6	662	VGNZCD	cell fusion glycop
13	66	26.6	662	S21382	cell fusion protei
14	65	26.2	631	VGNZPD	cell fusion glycop
15	65	26.2	631	A48346	cell fusion glycop
16	64	25.8	67	I78541	gonadoliberin prec
17	64	25.8	92	RHUG	gonadoliberin prec
18	63.5	25.6	98	I50739	gonadotropin-relea
19	62.5	25.2	89	I51423	gonadoliberin prec
20	60.5	24.4	82	I51365	gonadotropin-relea
21	60.5	24.4	555	B72486	probable hydantoin
22	60.5	24.4	880	B53743	protein-tyrosine k
23	60	24.2	546	S55386	cell fusion protei
24	59.5	24.0	90	RHWSG	gonadoliberin prec
25	59.5	24.0	1041	C83548	hypothetical prote
26	58.5	23.6	502	T36589	probable transmem
27	58	23.4	10	RHFGG	gonadoliberin - pi
28	58	23.4	10	RHSHG	gonadoliberin - sh
29	58	23.4	92	RHRTG	gonadoliberin prec

30 57.5 23.2 509 2 T45529
31 57.5 23.2 880 1 J04166
32 57 23.0 371 2 T49908
33 56.5 22.8 74 2 I51092
34 56.5 22.8 82 2 I51355
35 56.5 22.8 82 2 I51331
36 56 22.6 233 2 E87362
37 56 22.6 636 2 S47299
38 55.5 22.4 233 2 H69021
39 55.5 22.4 233 2 S38366
40 55 22.2 92 2 I50644
41 55 22.2 379 1 D8SPK
42 55 22.2 400 2 A10104
43 54.5 22.0 80 2 S39779
44 54.5 22.0 190 2 T37168
45 54.5 22.0 249 2 A41497

agaA protein (impo
protein-tyrosine k
hypothetical prote
gonadotropin relea
gonadotropin relea
gonadotropin relea
hypothetical prote
gene F protein - r
tetrahydromethanop
phosphoribosylamin
probable Galactosi
aldehyde reductase
probable tetr-fam
36K antigen pra -

ALIGNMENTS

RESULT 1

PQ0376
cell fusion glycoprotein - measles virus (strain TT) (fragment)
C:Species: measles virus
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C:Accession: PQ0376
R/Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A:Reference number: PQ0374; MUID:92300360; PMID:1607874
A:Accession: PQ0376
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.25; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

Qy 19 LSEIKGVIVHRLGV 33

Db 20 LSEIKGVIVHRLGV 34

RESULT 2

PQ0388
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)
C:Species: measles virus
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
R/Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A:Reference number: PQ0374; MUID:92300360; PMID:1607874
A:Accession: PQ0388
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.25; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

Qy 19 LSEIKGVIVHRLGV 33

Db 20 LSEIKGVIVHRLGV 34

RESULT 3

JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y
N/Contains: fusion glycoprotein F1; fusion glycoprotein F2
C/Species: subacute sclerosing panencephalitis virus, SSPEV
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C/Accession: JU0274

R;Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamahouchi, K.
Virus Genes 4, 173-181, 1990
A>Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A/Reference number: JU0274; MUID:90385702; PMID:1698327

A/Accession: JU0274
A/Molecule type: mRNA
A/Residues: 1-534 <KOM>
A/Cross-references: EMBL:DI0548; NID:G222256; PIDN:BA01405.1; PID:G222257
A/Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C/Genetics:

A/Gene: F
C/Superfamily: parainfluenza virus cell fusion protein
C/Keywords: Glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <PF2>
F;108-534/Product: cell fusion glycoprotein F1 #status predicted <PF1>
F;498-514/Domain: transmembrane #status predicted <TMN>
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

Db 288 LSEIKGVIVHRLGV 302
|||||

RESULT 4

S47300
Gene F protein - rinderpest virus
C/Species: rinderpest virus
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C/Accession: S47300; PQ0865
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL data library, March 1994

A/Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A/Reference number: S47299
A/Accession: S47300

A/Molecule type: DNA
A/Residues: 1-546 <EVA>
A/Cross-references: EMBL:Z31656; NID:G535406; PIDN:CA083482.1; PID:G535407
R;Chamberlain, R.W.; Wamwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J.
J. Gen. Virol. 74, 2775-2780, 1993

A>Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A/Reference number: PQ0865; MUID:94103786; PMID:8277286

A/Accession: PQ0865
A/Molecule type: mRNA
A/Residues: 86-191 <CHA>
C/Genetics:

A/Gene: F
C/Superfamily: parainfluenza virus cell fusion protein
C/Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 29.0%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

Db 284 LSEIKGVIVHRLGV 298
|||||

RESULT 5

C/Superfamily: parainfluenza virus cell fusion protein

E48556

cell fusion glycoprotein precursor - measles virus (strain AIK-C)

C/Species: measles virus
C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999

C/Accession: E48556
R;Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.

Virus Genes 7, 67-81, 1993

A>Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C

A/Reference number: A48556; MUID:93227570; PMID:8470368

A/Accession: E48556

A/Molecule type: genomic RNA

A/Residues: 1-550 <MOR>

A/Cross-references: GB:558435; NID:G299460; PIDN:AA26145.1; PID:G299465

A/Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)

C/Genetics:

A/Gene: F

C/Superfamily: parainfluenza virus cell fusion protein

C/Keywords: Glycoprotein; membrane fusion; transmembrane protein

F;1-22/Domain: signal sequence #status predicted <SIG>

F;23-107/Product: cell fusion glycoprotein F2 #status predicted <PF2>

F;108-550/Product: cell fusion glycoprotein F1 #status predicted <PF1>

F;113-138/Region: hydrophobic

F;495-514/Domain: transmembrane #status predicted <TMN>

F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 72; DB 1; Length 550;

Best Local Similarity 100.0%; Pred. No. 0.51;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

Db 288 LSEIKGVIVHRLGV 302
|||||

RESULT 6

VGNZMV

cell fusion glycoprotein precursor - measles virus

C/Species: measles virus

C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000

C/Accession: A26962; A25616; PQ0380; PQ0394

R;Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.

J. Gen. Virol. 68, 1695-1703, 1987

A>Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and comp

A/Reference number: A32794; MUID:87224816; PMID:3585281

A/Accession: A26962

A/Molecule type: mRNA

A/Residues: 1-553 <BUC>

A/Cross-references: GB:D00090; NID:G222061; PIDN:BA00056.1; PID:G222062

R;Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini,

Virol. 155, 508-523, 1986

A>Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles vir

A/Reference number: A94350; MUID:87071668; PMID:3788062

A/Accession: A25616

A/Molecule type: mRNA

A/Residues: 4-553 <RIC>

A/Cross-references: GB:ML4915; NID:G331762; PIDN:AAA46423.1; PID:G331763

A/Experimental source: strain Edmonston

R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A>Title: A measles virus isolate from a child with Kawasaki disease: sequence comparisc

A/Reference number: PQ0374; MUID:92300360; PMID:1607874

A/Accession: PQ0380

A/Molecule type: genomic RNA

A/Residues: 272-553 <SCH1>

A/Experimental source: isolate CL

A/Accession: PQ0384

A/Molecule type: genomic RNA

A/Residues: 272-553 <SCH2>

A/Experimental source: isolate SE

C/Genetics:

A/Gene: F

C;Keywords: glycoprotein; membrane fusion; transmembrane protein
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;16-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>
 F;111-553/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F;501-517/Domain: transmembrane #status predicted <TMN>
 F;32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 72; DB 1; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;

QY 19 LSEIKGVVHRLEGV 33
 DB 291 LSEIKGVVHRLEGV 305

RESULT 7
 VGNZRK
 cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
 F;1-19/Domain: signal sequence #status predicted <SIG>
 C;Species: rinderpest virus
 C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
 R;Hsu, D.; Yamahaka, M.; Miller, J.; Dale, B.; Grubman, M.; Vilma, T.
 Virology 166, 149-153, 1988
 A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis
 A;Reference number: A31051; PMID:88322864; PMID:3413983
 A;Accession: A31051
 A;Molecule type: genomic RNA
 A;Residues: 1-546 <HSU>
 C;Genetics:
 A;Gene: F
 C;Superfamily: parainfluenza virus cell fusion protein
 C;Keywords: glycoprotein; membrane fusion; transmembrane protein
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-108/Product: cell fusion glycoprotein F2 #status predicted <FF2>
 F;109-546/Product: cell fusion glycoprotein F1 #status predicted <FP2>
 F;109-134/Domain: transmembrane #status predicted <TMN>
 F;491-513/Domain: transmembrane #status predicted <TN2>
 F;25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.6%; Score 71; DB 1; Length 546;
 Best Local Similarity 93.3%; Pred. No. 0.68; Indels 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 0;

QY 19 LSEIKGVVHRLEGV 33
 DB 284 LSEIKGVVHRLEGV 298

RESULT 8
 S47305
 gene F protein - rinderpest virus
 C;Species: rinderpest virus
 C;Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
 C;Accession: S47305; S47301
 R;Baron, M.D.; Barrett, T.
 submitted to the EMBL Data Library, March 1994
 A;Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30
 A;Reference number: S47283
 A;Accession: S47305
 A;Molecule type: mRNA
 A;Residues: 1-546 <BAR>
 A;Cross-references: EMBL:Z30697; NID:G535396; PIDN:CAA83181.1; PID:G535401; EMBL:Z30700;
 C;Superfamily: parainfluenza virus cell fusion protein
 C;Keywords: transmembrane protein

Query Match 28.6%; Score 71; DB 2; Length 546;
 Best Local Similarity 93.3%; Pred. No. 0.68; Indels 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 0;

QY 19 LSEIKGVVHRLEGV 33

Db 284 LSEIKGVVHRLEGV 298

RESULT 9
 S47034
 cell fusion protein precursor - porpoise morbillivirus
 N;Alternate names: F protein
 C;Species: porpoise morbillivirus
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
 C;Accession: S47034
 R;Bolt, G.; Gottschalk, E.; Blixenkron-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.;
 submitted to the EMBL Data Library, July 1994
 A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbil
 A;Reference number: S47034
 A;Accession: S47034
 A;Molecule type: mRNA
 A;Residues: 1-552 <BO>
 A;Cross-references: EMBL:X80757; NID:G520639; PIDN:CAA56731.1; PID:G520640
 A;Experimental source: isolate Ulster 88
 A;Note: the source is designated as Cetacean morbillivirus
 C;Superfamily: parainfluenza virus cell fusion protein
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-552/Product: fusion protein #status predicted <MAT>

Query Match 26.8%; Score 66.5; DB 2; Length 552;
 Best Local Similarity 61.5%; Pred. No. 2.5; Indels 7; Gaps 1;
 Matches 16; Conservative 1; Mismatches 2;

QY 19 LSEIKGVVHRLEGVGSPSLHWSYGL 44
 DB 290 LSEIKGVVHRLEAV-----SYNL 308

RESULT 10
 JQ2223
 cell fusion protein F0 precursor - phocine distemper virus
 N;Contains: F1 and F2 chains
 C;Species: phocine distemper virus
 C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
 C;Accession: JQ2223
 R;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oerve
 J. Gen. Virol. 74, 1989-1994, 1993
 A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites
 e virus entity.
 A;Reference number: JQ2223; MUID:93389459; PMID:8376973
 A;Accession: JQ2223
 A;Molecule type: mRNA
 A;Residues: 1-542 <VIS>
 A;Cross-references: GB:L07075
 A;Note: the authors translated the codon ATC for residue 4 as Leu
 C;Comment: this fusion protein F0 is cleaved into F1 and F2 chains.
 C;Genetics:
 A;Gene: F
 C;Superfamily: parainfluenza virus cell fusion protein
 C;Keywords: glycoprotein; membrane fusion; transmembrane protein
 F;1-15/Domain: signal sequence #status predicted <SIG>
 F;16-542/Product: fusion protein #status predicted <MAT>
 F;16-99/Product: F2 chain #status predicted <F2C>
 F;105-542/Product: F1 chain #status predicted <F1C>
 F;105-135/Region: hydrophobic
 F;486-512/Domain: transmembrane #status predicted <TMW>
 F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 2; Length 542;
 Best Local Similarity 73.7%; Pred. No. 2.8; Indels 4; Gaps 0;
 Matches 14; Conservative 1; Mismatches 1;

QY 15 SLKLLSEIKGVVHRLEGV 33
 DB 276 SYPTLSEIKGVVHRLEAV 294

RESULT 11

Wed Mar 10 10:34:25 2004

us-09-848-834a-17.open.rpr

C:Accession: S21382
R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drillien, R.
Submitted to the EMBL Data Library, April 1992
A:Description: Vaccination of mice against canine distemper virus induced encephalitis
C:Species: rinderpest virus
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C:Accession: A28921
R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the
A:Reference number: A28921; MUID:88219541; PMID:3285575
A:Accession: A28921
A:Molecule type: mRNA
A:Residues: 1-546 <TSU>
A:Cross-references: GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F:109-133/Domain: transmembrane #status predicted <TM1>
F:1485-513/Domain: transmembrane #status predicted <TM2>
F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 2.8;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 284 LSEIKGVIVHRLGV 298
|||||

RESULT 12
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: canine distemper virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: J50321
R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A:Reference number: J50321; MUID:88129050; PMID:3433924
A:Accession: J50321
A:Molecule type: mRNA
A:Residues: 1-662 <BAR>
A:Cross-references: GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-135/Domain: signal sequence #status predicted <SIG>
F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F:606-629/Domain: transmembrane #status predicted <MEM>
F:62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 3.5;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLSSEIKGVIVHRLGV 33
DB 396 SYPTLSEVKGIVHRLGV 414
|||||

RESULT 13
S21382
cell fusion protein - canine distemper virus
C:Species: canine distemper virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999

C:Accession: S21382
R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drillien, R.
Submitted to the EMBL Data Library, April 1992
A:Description: Vaccination of mice against canine distemper virus induced encephalitis
C:Species: rinderpest virus
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C:Accession: A28921
R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the
A:Reference number: A28921; MUID:88219541; PMID:3285575
A:Accession: A28921
A:Molecule type: mRNA
A:Residues: 1-546 <TSU>
A:Cross-references: GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F:109-133/Domain: transmembrane #status predicted <TM1>
F:1485-513/Domain: transmembrane #status predicted <TM2>
F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 2; Length 662;
Best Local Similarity 73.7%; Pred. No. 3.5;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLSSEIKGVIVHRLGV 33
DB 396 SYPTLSEVKGIVHRLGV 414
|||||

RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C:Accession: JQ1368
R:Koevamees, J.; Slixenkron-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglutinin
A:Reference number: JQ1368; MUID:92113538; PMID:1765768
A:Accession: JQ1368
A:Molecule type: genomic RNA
A:Residues: 1-631 <KOV>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:189-106/Domain: transmembrane #status predicted <TM1>
F:189-193/Region: cleavage processing #status predicted
F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F:194-212/Domain: transmembrane #status predicted <TM2>
F:575-595/Domain: transmembrane #status predicted <TM3>
F:110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 4.4;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLSSEIKGVIVHRLGV 33
DB 365 SYPTLSEVKGIVHRLGV 383
|||||

RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C:Accession: A48346
R:Curran, M.D.; Lu, Y.-J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced ami
A:Reference number: A48346; MUID:92398437; PMID:1524494
A:Accession: A48346
A:Molecule type: mRNA
A:Residues: 1-631 <CUR>
A:Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBI:P:113099)
C:Genetics:
A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein
C1:188/Domain: transmembrane #status predicted <P2>
C1:189/Domain: cell fusion glycoprotein f2 #status predicted <P2>
C1:190/Domain: transmembrane #status predicted <TM1>
C1:191-196/Product: cell fusion glycoprotein f1 #status predicted <P1>
C1:192-631/Product: transmembrane #status predicted <TM2>
C1:194-219/Domain: transmembrane #status predicted <TM3>
C1:575-595/Domain: transmembrane #status predicted <TM3>
C1:595-610/Binding site carbohydrate (Asn) (covalent) #status predicted <P1>
C1:610-142/Binding site carbohydrate (Asn) (covalent) #status predicted <P1>

Query Match 26.2%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 4.4;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 15 SLKLLSEIKGVIVHRLGV 33
pB 365 SYPTLSEIKGVVHRLGV 383

Search completed: March 10, 2004, 09:16:51
Job time : 14.6892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 8.59533 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSGSPSLKLLS.....HRLEGVEGSLHWSYGLRXP 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72.5	29.2	95	1 GONL SPAU	P51919 sparus aura
2	72	29.0	534	1 VGLF MEAS	P25032 measles vir
3	72	29.0	546	1 VGLF RIND	P41360 rinderpest
4	72	29.0	550	1 VGLF MEAS	P35973 measles vir
5	72	29.0	550	1 VGLF MEAS	P08300 measles vir
6	71	28.6	546	1 VGLF RIND	P41356 rinderpest
7	68.5	27.6	95	1 GONL PAGMA	P70074 pagrus majo
8	66	26.6	546	1 VGLF RIND	P18864 rinderpest
9	66	26.6	662	1 VGLF CDVO	P13569 canine dist
10	65	26.2	546	1 VGLF RIND	P12574 rinderpest
11	65	26.2	631	1 VGLF PHODV	P28886 phocine dis
12	64	25.8	67	1 GONL MACMU	P55247 macaca mula
13	64	25.8	92	1 GONL HUMAN	P01148 homo sapien
14	64	25.8	529	1 VGLF MEAS	P26031 measles vir
15	63.5	25.6	94	1 GONL HAPBU	P51918 haplochromi
16	62.5	25.2	89	1 GONL XENLA	P45656 xenopus lae
17	61.5	24.8	61	1 GONL SHEEP	P25888 ovis aries
18	60.5	24.4	74	1 GON3 ONCMY	P55246 oncorhynch
19	60.5	24.4	82	1 GON3 SALTR	P45653 salmo trutt
20	60.5	24.4	169	1 CX41 THUOB	Q918u0 thunnus obe
21	60.5	24.4	880	1 TY03 MUOSE	P51144 mus musculu
22	60	24.2	92	1 GONL TURGE	Q95335 tupaia glis
23	59.5	24.0	90	1 GONL MOUSE	P13562 mus musculu
24	59.5	24.0	91	1 GONL PIG	P19921 sus scrofa
25	59	23.8	91	1 GONL MORSA	O73812 morone saxa
26	58	23.4	63	1 GON1 MESSAU	O09163 mesocricetu
27	58	23.4	90	1 GON1 RANCA	P09763 rattus catesb
28	58	23.4	92	1 GON1 RAT	P07490 rattus norv
29	58	23.4	99	1 GON1 DICCLA	Q91a10 dicerorarch
30	57.5	23.2	880	1 TY03 RAT	P55146 rattus norv
31	56.5	22.8	74	1 GON3 ONCTS	Q92097 oncorhynch
32	56.5	22.8	82	1 GON3 SALSA	P35629 salmo salar
33	55.5	22.4	233	1 MTRD METHH	O27230 methanobact

34 55.5 22.4 233 1 MTRD METTM P80183 methanobact
35 55 22.2 92 1 GON1 CHICK P37042 gallus gall
36 55 22.2 379 1 PURK BACSU P12045 bacillus su
37 54.5 22.0 249 1 PRA MYCLE P41484 mycobacteri
38 54 21.8 10 1 GON1 ALLMI P37041 alligator m
39 54 21.8 465 1 STHA_SALTY Q8xfp5 salmonella
40 53 21.4 497 1 DLHD_WANSE O18480 manduca sex
41 53 21.4 508 1 VGLG_THNV P09223 infectious
42 52.5 21.2 217 1 RS3_BACST P23309 bacillus at
43 52.5 21.2 481 1 DOK1_HUMAN Q99704 homo sapien
44 52 21.0 91 1 GON1_ORYLA Q9dgc8 o prognado
45 52 21.0 132 1 EAT2_HUMAN O14796 homo sapien

ALIGNMENTS

RESULT 1
GONL SPAU STANDARD; PRT; 95 AA.
AC
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
(LH-RH I) (Luliberin I) (SBGnRH).
GN GNRH1.
OS Sparus aurata (Gilthead sea bream)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Necteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
[1]
RN SEQUENCE FROM N.A.
RP
RC TISSUE=Brain;
RX MEDLINE=95268499; PubMed=7749463;
RA Gothif Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
releasing hormone from the gilthead seabream (Sparus aurata).";
EL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE=Brain;
RX MEDLINE=95083645; PubMed=7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RT Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
"Three forms of gonadotropin-releasing hormone characterized from
brains of one species.";
EL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
RP
RL
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U30320; AAA75469.1;
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Signal; Multigene family; Pyrrolidone carboxylic acid.
FT SIGNAL 1 25
FT CHAIN 26 95 PRONADOLIBERIN I.

FT PEPTIDE 26 35 GONADOLIBERIN I.
 FT PEPTIDE 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
 FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FD6B87DA CRC64;

Query Match 29.2%; DB 1; Length 95;
 Best Local Similarity 45.7%; Pred. No. 0.016; Mismatches 15; Indels 1; Gaps 1;
 Matches 16; Conservative 3;

QY 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLGV 36
 DB 27 HWSYGLRPGSGK-RDLDSLSLTGLNLTIERFPHVDSF 60

RESULT 2
 VGLF_MEASV STANDARD; PRT; 534 AA.
 AC P26032;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fusion glycoprotein precursor (Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1).
 GN F.
 OS Measles virus (strain Yamagata-1) (Subacute sclerosing panencephalitis virus).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 OX NCBI_TaxID=11239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90385702; PubMed=1698327;
 RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;
 RT "Molecular analysis of structural protein genes of the Yamagata-1 strain of defective subacute sclerosing panencephalitis virus. IV. Nucleotide sequence of the fusion gene."
 RL Virus Genes 4:173-181(1990).
 CC -!- FUNCTION: This protein directs fusion of viral and cellular membranes.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
 CC -!- LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10548; BAA01405.1; -
 DR HSSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion gly.
 DR Pfam; PF00523; Fusion gly; 1.
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 534 FUSION GLYCOPROTEIN F0.
 FT CHAIN 24 112 PROTEIN F2.
 FT CHAIN 113 534 PROTEIN F1.
 FT TRANSMEM 113 136 POTENTIAL.
 FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 495 515 POTENTIAL.
 FT DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 58 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;

Query Match 29.0%; Score 72; DB 1; Length 534;
 Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;
 Matches 15; Conservative 0;

Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;
 Matches 15; Conservative 0;

QY 19 LSEIKGVIVHRLGV 33
 DB 288 LSEIKGVIVHRLGV 302

RESULT 3
 VGLF_RINDB STANDARD; PRT; 546 AA.
 AC P41360;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fusion glycoprotein precursor (Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1).
 GN F.
 OS Rinderpest virus (strain RB1) (RDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 OX NCBI_TaxID=39007;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95086609; PubMed=7996154;
 RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
 RT "Nucleotide sequence comparisons of the fusion protein gene from virulent and attenuated strains of rinderpest virus."
 RL J. Gen. Virol. 75:3611-3617(1994).
 CC -!- FUNCTION: This protein directs fusion of viral and cellular membranes.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
 CC -!- LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z31656; CAA83482.1; -
 DR PIR; S47300; S47300.
 DR HSSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion gly.
 DR Pfam; PF00523; Fusion gly; 1.
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
 FT CHAIN 20 108 F2 PROTEIN.
 FT CHAIN 109 546 F1 PROTEIN.
 FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
 FT TRANSMEM 109 133 POTENTIAL.
 FT TRANSMEM 484 513 POTENTIAL.
 FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
 FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
 SQ SEQUENCE 546 AA; 58418 MW; 3B539B989344F401 CRC64;

Query Match 29.0%; Score 72; DB 1; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;
 Matches 15; Conservative 0;

QY 19 LSEIKGVIVHRLGV 33
 DB 284 LSEIKGVIVHRLGV 298

```

RESULT 4
VGLF MEASA
ID _VGLF MEASA STANDARD; PRT; 550 AA.
AC P35973.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain A1K-C) (Subacute sclerosing panencephalitis
OS virus).
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=36408;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93227570; PubMed=8470368;
RA Mori T., Sasaki K., Hashimoto H., Makino S.;
RT "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT the A1K-C strain of attenuated measles virus.";
RL Virus Genes 7:67-81(1993).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC -!- LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S58435; AAB26145.1; -.
DR PIR; E48556; E48556.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 POTENTIAL F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59540 MW; AAC4DBA92DE0D938 CRC64;

Query Match 29.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 288 LSEIKGVIVHRLGV 302

RESULT 5
VGLF MEASE
ID _VGLF MEASE STANDARD; PRT; 550 AA.
AC P08300.
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Edmonston) (Subacute sclerosing panencephalitis
OS virus);
OS Measles virus (strain Halle) (Subacute sclerosing panencephalitis
OS virus);
OS Measles virus (strain Leningrad-16) (Subacute sclerosing panencephalitis
OS virus);
OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerosing
OS panencephalitis virus);
OS Measles virus (strain Philadelphia-26) (Subacute sclerosing
OS panencephalitis virus); and
OS Measles virus (strain Edmonston B) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=Edmonston;
RX MEDLINE=87071668; PubMed=3788062;
RA Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
RA Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT "The nucleotide sequence of the mRNA encoding the fusion protein of
RT measles virus (Edmonston strain): a comparison of fusion proteins
RT from several different paramyxoviruses.";
RL Virology 155:508-523(1986).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=Halle;
RX MEDLINE=87224816; PubMed=3585281;
RA Buckland R., Gerald C., Barker R., Wild T.F.;
RT "Fusion glycoprotein of measles virus: nucleotide sequence of the
RT gene and comparison with other paramyxoviruses.";
RL J. Gen. Virol. 68:1695-1703(1987).
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=Edmonston;
RX MEDLINE=90085790; PubMed=2596022;
RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,
RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT "Mutated and hypermutated genes of persistent measles viruses which
RT caused lethal human brain diseases.";
RL Virology 173:415-425(1989).
RN [4]
RN SEQUENCE FROM N.A.
RX STRAIN=Edmonston;
RX MEDLINE=92263801; PubMed=1585658;
RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
RA Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus.";
RL Virology 188:910-915(1992).
RN [5]
RN SEQUENCE FROM N.A.
RX STRAIN=Edmonston, Leningrad-16, and Edmonston-Zagreb;
RX MEDLINE=94249283; PubMed=8191786;
RA Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT "Comparison of sequences of the H, F, and N coding genes of measles
RT virus vaccine strains.";
RL Virus Res. 31:317-330(1994).
RN [6]
RN SEQUENCE FROM N.A.
RX STRAIN=Philadelphia-26;
RX MEDLINE=94303181; PubMed=8030232;
RA Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT "Restriction of fusion protein mRNA as a mechanism of measles virus
RT persistence.";
RL Virology 202:665-672(1994).
RN [7]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston B;
RA Billster M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE NATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14915; AAA46423.1; -
DR EMBL; X05597; CAA29090.1; ALT_INIT.
DR EMBL; K01711; AAA75498.1; ALT_INIT.
DR EMBL; K01711; AAA75499.1; -
DR EMBL; U03657; AAA56647.1; -
DR EMBL; U03659; AAA56648.1; ALT_INIT.
DR EMBL; U03670; AAA56660.1; ALT_INIT.
DR EMBL; U08415; AAA50550.1; ALT_INIT.
DR EMBL; Z66517; CAA91367.1; ALT_INIT.
DR EMBL; Z66517; CAA91368.1; -
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 113 550 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59532 MW; 7AA4F1CA82169093 CRC64;

Query Match 29.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Gaps 0;
Matches 15; Conservative 0; Indels 0;

Qy 19 LSEIKGVVHRLEGV 33
Db 288 LSEIKGVVHRLEGV 302

RESULT 6
VGLF_RINDR STANDARD; PRT; 546 AA.
AC P41356;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RBOK) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36409;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95088609; PubMed=7996154;
RX

```

```

RA Evans S.A.; Baron M.D.; Chamberlain R.W.; Goatley L.; Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus."
RL J. Gen. Virol. 75:3611-3617(1994).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE NATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CAA83186.1; -
DR EMBL; Z30697; CAA83181.1; -
DR PR; S47305; S47305.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 109 546 F2 PROTEIN.
FT CHAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DF8AFFDBECB95 CRC64;

Query Match 28.6%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.17; Mismatches 1; Gaps 0;
Matches 14; Conservative 1; Indels 0;

Qy 19 LSEIKGVVHRLEGV 33
Db 284 LSEIKGVVHRLEGV 298

RESULT 7
GONI_PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Okuzawa K.; Granneman J.; Bogerd J.; Goos H.; Zohar Y.; Kagawa H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC similarity).

```

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GHRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D86582; BAA13129.1; -
CC InterPro; IPR002012; GHRH.
CC InterPro; IPR004079; Gonadoliberin.
CC Pfam; PF00446; GHRH.1.
CC PRINTS; PRO1541; GONADOLIBERN.
CC PROSITE; PS00473; GHRH.1.
CC Cleavage on pair of basic residues; Hormones; Amidation; Hypothalamus;
KW Signal; Multigene family; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 95 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 95 GHRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;
Query Match 27.6%; Score 58.5; DB 1; Length 95;
Best Local Similarity 42.9%; Pred. No. 0.054;
Matches 15; Conservative 3; Mismatches 16; Indels 1; Gaps 1;
QY 2 HWSVGLRGPSPKLLSEIKGVVHRLEGGP 36
DB 25 HWSVGLSPGGR-RLLDSLSLDLGIIEFFPHADSP 58
RESULT 8
VGLF_RINDL
ID VGLF_RINDL STANDARD; PRT; 546 AA.
AC P10864;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
OS F.
OS Rinderpest virus (strain L) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11243;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219541; PubMed=3285575;
RA Tsukiyama K., Yoshikawa Y., Yamanouchi K.;
RT "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT sequence of the F mRNA, and several features of the F protein.";
RL Virology 164:523-530(1988).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC EMBL; M20870; AAA47399.1; -
CC PIR; A28921; VGNZEL.
CC HSP; P04849; ISVF.
CC InterPro; IPR000776; Fusion.gly.
CC Pfam; PF00523; fusion.gly.1.
KW Glycoprotein; Fusion Protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;
Query Match 26.6%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.78;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 19 LSEIKGVVHRLEGV 33
DB 284 LSEIKGVVHRLESV 298
RESULT 9
VGLF_CDVO
ID VGLF_CDVO STANDARD; PRT; 662 AA.
AC P12569; Q65991;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
OS F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88129050; PubMed=3433924;
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT "The nucleotide sequence of the gene encoding the F protein of canine
RT distemper virus: a comparison of the deduced amino acid sequence with
RT other paramyxoviruses".
RL Virus Res. 8:373-386(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227696; PubMed=8470428;
RA Wild T.P., Bernard A., Spehner D., Villevial D., Drillien R.;
RT "Vaccination of mice against canine distemper virus-induced
RT encephalitis with vaccinia virus recombinants encoding measles or
RT canine distemper virus antigens".
RL Vaccine 11:438-444(1993).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21849; AAA42878.1; -.
CC DR EMBL; X65509; CAA46481.1; -.
CC DR PIR; J80321; VGNZCD.
CC DR PIR; S21382; S21382.
CC DR HSSP; P04849; LSVP.
CC DR InterPro; IPR000776; Fusion gly.
CC DR Pfam; PF00523; fusion gly; 1.
CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC FT SIGNAL 1 662 FUSION GLYCOPROTEIN F0.
CC FT CHAIN ? 662 PROTEIN F2.
CC FT CHAIN ? 224 PROTEIN F1.
CC FT CHAIN 225 662 POTENTIAL.
CC FT TRANSMEM 606 629 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC FT DISULFID 180 307 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CONFLICT 3 3 R -> K (IN REF. 2).
CC FT CONFLICT 140 140 D -> N (IN REF. 2).
CC FT CONFLICT 152 152 N -> S (IN REF. 2).
CC FT CONFLICT 171 171 I -> M (IN REF. 2).
CC FT CONFLICT 174 174 L -> V (IN REF. 2).
CC FT CONFLICT 662 662 L -> H (IN REF. 2).
CC SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 26.6%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.96;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLSEIKGVIVHRLGV 33
DB 396 SYPTUSEVKGIVHRLGV 414

RESULT 10
VGLF RINDK STANDARD; PRT; 546 AA.
AC P12574;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain Kabete O) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11242;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88322864; PubMed=3413983;
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Vilma T.;
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT analysis with other morbilliviruses.";
RL Virology 166:149-153(1988).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; M21514; AAA47400.1; -.
DR PIR; A31051; VGNZK.
DR HSSP; P04849; LSVP.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; fusion gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 546 F2 PROTEIN.
FT CHAIN 20 108 F1 PROTEIN.
FT CHAIN 109 546 ARG-RICH (BASIC).
FT DOMAIN 104 108 POTENTIAL.
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 ARG/LYS-RICH (BASIC).
FT DOMAIN 514 517 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT DISULFID 64 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58662 MW; 476D74DCC18BCFCF CRC64;

Query Match 26.2%; Score 65; DB 1; Length 546;
Best Local Similarity 86.7%; Pred. No. 1;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 284 LSEIKGVIVHRLGV 298

RESULT 11
VGLF PHODV STANDARD; PRT; 631 AA.
AC P28836;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Phocine distemper virus (PDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11240;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate DK88-4A;
RX MEDLINE=92113538; PubMed=1765768;
RA Keenanees J., Blixenkron-Moeller M., Sharma B., Oervell C.,
RA Norby E.;
RT "The nucleotide sequence and deduced amino acid composition of the
RT haemagglutinin and fusion proteins of the morbillivirus phocid
RT distemper virus.";
RL J. Gen. Virol. 72:2959-2966(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ulster/88;
RX MEDLINE=92398437; PubMed=1524494;
RA Curran M.D., Lu Y.J., Rima B.K.;
RT "The fusion protein gene of phocine distemper virus: nucleotide and
RT deduced amino acid sequences and a comparison of morbillivirus fusion
RT proteins.";
RL Arch. Virol. 126:159-169(1992).
RN [3]
RP SEQUENCE OF 95-631 FROM N.A.
RX STRAIN=Ulster/88;
RX MEDLINE=91089508; PubMed=2264246;
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;
RT "Nucleotide sequence analysis of phocine distemper virus reveals its
RT distinctness from canine distemper virus.";
RL Vet. Rec. 127:430-431(1990).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.

```


CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
 CC LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10371; BAA01206.1; --
 CC FIRM: A48346; A48346.
 CC PIR: JQ1368; VGNZPD.
 CC HSSP: P04849; LSVE.
 CC InterPro: IPR000776; Fusion gly.
 CC Pfam: PF00523; fusion gly; 1.
 CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
 FT SIGNAL 1 631 FUSION GLYCOPROTEIN P0.
 FT CHAIN ? 188 F2 PROTEIN.
 FT CHAIN 194 631 F1 PROTEIN.
 FT CHAIN 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 FT DISULFID 149 276 POTENTIAL.
 FT TRANSMEM 89 106 POTENTIAL.
 FT TRANSMEM 194 212 POTENTIAL.
 FT TRANSMEM 575 595 POTENTIAL.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 63 63 I -> V (IN REF. 2).
 SQ SEQUENCE 631 AA; 68873 MW; D1PC87CDD426E9B8 CRC64;
 Query Match 26.2%; Score 65; DB 1; Length 631;
 Best Local Similarity 68.4%; Pred. No. 1.2; Length 631;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 15 SLKLLSEIKGVIVHRLEGV 33
 DB 365 SVFTLSEVKGVVHSLV 383
 RESULT 12
 GONI_MACMU STANDARD; PRT; 67 AA.
 AC P55247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
 DE (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=95124501; PubMed=7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "Developmental expression of the genes encoding transforming growth
 RT factor alpha and its receptor in the hypothalamus of female rhesus
 RT macaques."
 RL Neuroendocrinology 60:346-359(1994).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the GnRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S75918; AAB33096.1; --
 CC PIR: I78541; I78541.
 CC InterPro: IPR002012; GnRH.
 CC InterPro: IPR004079; Gonadolibirin I.
 CC Pfam: PF00446; GnRH; 1.
 CC PRINTS: PR01541; GONADOLIBIRNI.
 CC PROSITE: PS00473; GnRH; 1.
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1
 FT SIGNAL 1 5 BY SIMILARITY.
 FT CHAIN 6 >67 PROGNADOLIBIRIN I.
 FT PEPTIDE 6 15 GONADOLIBIRIN I.
 FT PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACT_SITE 8 8 ACTIVITY (BY SIMILARITY).
 FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 15 15 SIMILARITY).
 FT MOD_RES 67 67 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
 FT MOD_RES 67 67 SIMILARITY).
 FT NON_TER 67 67
 SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
 Query Match 25.8%; Score 64; DB 1; Length 67;
 Best Local Similarity 40.0%; Pred. No. 0.14;
 Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 2 HWSYGLRPGSGPSLKLSEIKGVI 26
 DB 7 HWSYGLRPGGKRDAENLMDSFQEI 31
 RESULT 13
 GONI_HUMAN STANDARD; PRT; 92 AA.
 AC P01138;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
 DE peptide I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene."
 RN Nucleic Acids Res. 17:6403-6403(1989).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).

SQ SEQUENCE 529 AA; 57331 MW; AE987BC9F07E9AA9 CRC64;

Query Match 25.8%; Score 64; DB 1; Length 529;
 Best Local Similarity 93.3%; Pred. No. 1.4;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 |||||
 DB 291 LSEIKGVIVHRLEGV 305

RESULT 15

GONI_HAPBU STANDARD; PRT; 94 AA.

AC PS1918; O93387;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Progonadoliberein I precursor (Contains: Gonadoliberein I (Luteinizing hormone releasing hormone I) (Gonadotropin-releasing hormone I))
 DE (GNRH-I) (h-RH I) (Inhiberin I); GNRH-associated peptide I].
 GN GNRH.

OS Haplochromis burtoni (Burton's mouthbrooder).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Perciformes; Labroidae;
 OC Cichlidae; Astatotilapia.
 OX NCBI_taxid=8153;
 RN [1]

SEQUENCE FROM N.A.
 RP MEDLINE=95396797; PubMed=7667296;
 RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
 RT "Three gonadotropin-releasing hormone genes in one organism suggest novel roles for an ancient peptide";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367 (1995).
 RN [2]

SEQUENCE FROM N.A.
 RP MEDLINE=99061842; PubMed=9843638;
 RA White R.B., Fernald R.D.;
 RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression reveals a distinct origin for GNRH-containing neurons in the midbrain";
 RL Gen. Comp. Endocrinol. 112:322-329 (1998).
 RN [3]

SEQUENCE OF 23-32, AND MASS SPECTROMETRY.
 RP TISSUE=Pituitary;
 RX MEDLINE=95372591; PubMed=7644702;
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,
 RA Sherwood N.M.;
 RT "Primary structure of solitary form of gonadotropin-releasing hormone (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid and pumpkinseed fish";
 RL Regul. Pept. 57:43-53 (1995).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-GONADAL AXIS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPHYSAL AXONS.
 CC -!- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.
 CC -!- SIMILARITY: Belongs to the GNRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U31865; AAC59691.1; -.
 CC EMBL: AF076961; AAC27716.1; -.

DR PIR; I50739; I50739.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
 DR GO; GO:0007275; P:development; IDA.
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadolibereinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 22
 FT CHAIN 23 94
 FT PEPTIDE 23 32
 FT PEPTIDE 36 94
 FT MOD_RES 23 23
 FT MOD_RES 32 32
 FT CONFLICT 86 94
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA83333278D7 CRC64;
 25.6%; Score 63.5; DB 1; Length 94;
 Best Local Similarity 42.9%; Pred. No. 0.24; 17; Indels 1; Gaps 1;
 Matches 15; Conservative 2; Mismatches 2

QY 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLEGV 36
 |||||
 DB 24 HWSYGLSPGSK-RDLDFNSDTLGNWVEFPV 57

Search completed: March 10, 2004, 09:13:58
 Job time : 9.59533 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 46.4514 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-17
Perfect score: 248
Sequence: 1 XHWSYGLRPGSSPSLKLLS.....HRLGVGSPSLHWSYGLRFX 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp Vertebrate:*
 - 14: sp Unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72.5	29.2	87	13 Q9Y126	Q9Y126 sparus aura
2	72	29.0	534	12 Q04243	Q04243 measles vir
3	72	29.0	537	12 Q04242	Q04242 measles vir
4	72	29.0	545	12 Q9PXA4	Q9PXA4 measles vir
5	72	29.0	546	12 Q9LHA5	Q9LHA5 rinderpest
6	72	29.0	550	12 P90331	P90331 measles vir
7	72	29.0	550	12 Q9QX00	Q9QX00 measles vir
8	72	29.0	550	12 Q9QW99	Q9QW99 measles vir
9	72	29.0	550	12 P90330	P90330 measles vir
10	72	29.0	550	12 Q9QW77	Q9QW77 measles vir
11	72	29.0	550	12 Q9WMK4	Q9WMK4 measles vir
12	72	29.0	550	12 Q89495	Q89495 measles vir
13	72	29.0	550	12 Q8V049	Q8V049 measles vir
14	72	29.0	550	12 Q9YJ94	Q9YJ94 measles vir
15	72	29.0	550	12 Q9QEX1	Q9QEX1 measles vir
16	72	29.0	550	12 Q9QEW8	Q9QEW8 measles vir

17	72	29.0	553	12 Q93055	Q93055 measles vir
18	72	29.0	553	12 Q9IC36	Q9IC36 measles vir
19	72	29.0	553	12 P88973	P88973 measles vir
20	72	29.0	553	12 Q83536	Q83536 measles vir
21	72	29.0	553	12 Q11383	Q11383 measles vir
22	72	29.0	553	12 Q91FK2	Q91FK2 measles vir
23	72	29.0	553	12 Q83533	Q83533 measles vir
24	72	29.0	553	12 Q83525	Q83525 measles vir
25	72	29.0	553	12 Q83518	Q83518 measles vir
26	72	29.0	553	12 P88974	P88974 measles vir
27	72	29.0	553	12 Q83527	Q83527 measles vir
28	72	29.0	553	12 Q83521	Q83521 measles vir
29	72	29.0	553	12 Q83530	Q83530 measles vir
30	72	29.0	553	12 Q91248	Q91248 measles vir
31	72	29.0	553	12 Q91Q22	Q91Q22 measles vir
32	72	29.0	553	12 Q04244	Q04244 measles vir
33	72	29.0	579	12 Q9PMU4	Q9PMU4 measles vir
34	71.5	28.8	68	13 Q8JIF4	Q8JIF4 acanthopagr
35	68.5	27.6	64	13 Q8JIF3	Q8JIF3 dentex dent
36	68.5	27.6	64	13 Q8JIF2	Q8JIF2 pagrus majo
37	68	27.4	545	12 Q9QEW6	Q9QEW6 measles vir
38	68	27.4	553	12 Q11380	Q11380 measles vir
39	67.5	27.2	96	13 Q8UW80	Q8UW80 verasper mo
40	66.5	26.8	552	12 Q66147	Q66147 cetacean mo
41	66	26.8	528	12 Q9YJW9	Q9YJW9 canine dist
42	66	26.6	530	12 Q8QV06	Q8QV06 canine dist
43	66	26.6	662	12 Q9DXZ2	Q9DXZ2 canine dist
44	66	26.6	662	12 Q91KN3	Q91KN3 canine dist
45	66	26.6	662	12 Q9YKL7	Q9YKL7 canine dist

ALIGNMENTS

RESULT 1

Q9Y126 PRELIMINARY; PRT; 87 AA.

AC Q9Y126; TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 24, Last annotation update)

DE SBGNRH (Gonadoliberin) (Gonadotropin-releasing hormone) (LH-RH)

DE (Luliberin) (Fragment).

OS Sparus aurata (Gilthead sea bream).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;

OC Sparidae; Sparus.

OX [1] NCBI_TaxID=8175;

RN [1] SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RA Nabissi M.;

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

DR EMBL; AF046801; AAD02427.1; .

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.

DR GO; GO:0007275; P:development; IEA.

DR InterPro; IPR002012; GNRH.

DR InterPro; IPR004079; GonadoliberinI.

DR Pfam; PF00446; GNRH; 1.

DR PRINTS; PR01541; GONADOLIBRNI.

DR	PROSITE; PS00473; GNRH; 1.
DR	Amidation; Hormone.
FT	NON_TER 87
FT	NON_TER 87
SQ	SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;
Query Match 29.2%; Score 72.5; DB 13; Length 87;	
Best Local Similarity 45.7%; Pred. No. 0.12;	

RA	Cattaneo R., Billetter M.A.;
RL	Virology 0:0-0(0);
RM	EMBL; X16567; CAA34574.1; -;
DR	EMBL; X16567; CAA34575.1; -;
DR	HSSP; P04849; ISVF.
DR	GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR	GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR	InterPro; IPR000776; Fusion.gly.
DR	Pfam; PF00523; fusion.gly; I.
DR	SEQUENCE 537 AA; 58275 MW; D0A60AC6D979E06 CRC64;
SQ	
 Query Match 29.0%; Score 72; DB 12; Length 537; Best Local Similarity 100.0%; Pred.No. 1.1; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	19 LSEIKGVIVHLEGV 33
DB	291 LSEIKGVIVHLEGV 305
 RESULT 4 Q9PXA4 ID Q9PXA4 PRELIMINARY; PRT; 545 AA. AC Q9PXA4; DT 01-MAY-2000 (TREMBLrel. 13, Created) DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update) DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update) DE Fusion protein. OS Measles virus. OC Viruses; ssRNA negative-strand viruses; Mononegavirales; OC Paramyxoviridae; Paramyxovirinae; Morbillivirus. OX NCBI_TaxID=11234; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=OSA-3; RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M., RA Ogura H.; RT "Nucleotide sequences of the fusion protein gene of subacute RT sclerosing panencephalitis viruses; deduced amino acid sequences RT showed the cytoplasmic domain highly mutated --truncated, elongated or RT predicted secondary structure changed."; RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. DR EMBL; AF179440; AAF02705.1; -; DR EMBL; AF179439; AAF02704.1; -; DR HSSP; P04849; ISVF. DR GO; GO:0019039; F:viral-cell fusion molecule activity; IEA. DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA. DR InterPro; IPR000776; Fusion.gly. DR Pfam; PF00523; fusion.gly; I. DR SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;	
 Query Match 29.0%; Score 72; DB 12; Length 545; Best Local Similarity 100.0%; Pred.No. 1.1; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	19 LSEIKGVIVHLEGV 33
DB	288 LSEIKGVIVHLEGV 302
 RESULT 5 Q91HA5 ID Q91HA5 PRELIMINARY; PRT; 546 AA. AC Q91HA5; DT 01-DEC-2001 (TREMBLrel. 19, Created) DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update) DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update) DE Fusion protein. OS Rinderpest virus. OC Viruses; ssRNA negative-strand viruses; Mononegavirales; OC Paramyxoviridae; Paramyxovirinae; Morbillivirus. OX NCBI_TaxID=11241;	

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=K.
RX  MEDLINE=2104265; PubMed=11186456;
RA  Alanot P.K., Smelev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA  Gusev A.A.;
RT  "Primary structure of the F-gene from Rinderpest virus strain K.";
RL  Mol. Gen. Microbiol. Virusol. 4:29-33(2000).
RN  SEQUENCE FROM N.A.
RP  STRAIN=K.
RC  Ayanot P.K., Smelev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA  Gusev A.A.;
RL  Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY035887; AAK63190.1; -.
DR  PIR; PQ0866; PQ0866.
DR  PIR; PQ0867; PQ0867.
DR  PIR; PQ0873; PQ0873.
DR  GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR  GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR  InterPro; IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly.1.
SQ  SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F0B CRC64;

Query Match          29.0%; Score 72; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  19 LSEIKGVIVHRLEGV 33
DB  284 LSEIKGVIVHRLEGV 298

RESULT 6
P90331 ID P90331 PRELIMINARY; PRT; 550 AA.
AC P90331;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
EN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
EN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63926; BAA09958.1; -.
DR EMBL; AF179431; AAF02696.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04949; 1SVF.

```

```

DR  GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR  GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR  InterPro; IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly.1.
SQ  SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match          29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  19 LSEIKGVIVHRLEGV 33
DB  288 LSEIKGVIVHRLEGV 302

RESULT 7
Q9QEX0 ID Q9QEX0 PRELIMINARY; PRT; 550 AA.
AC Q9QEX0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Toyoshima;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR  GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR  GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR  InterPro; IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly.1.
SQ  SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match          29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  19 LSEIKGVIVHRLEGV 33
DB  288 LSEIKGVIVHRLEGV 302

RESULT 8
Q9QEW9 ID Q9QEW9 PRELIMINARY; PRT; 550 AA.
AC Q9QEW9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute

```

```

RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AFI19436; AAF02701.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0A56DBFC5DD228BA CRC64;

Query Match      29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302

RESULT 9
ID P90330 PRELIMINARY; PRT; 550 AA.
AC P90330;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=Nagahata (HB);
RA Sheng J., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63924; BAA03951.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match      29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302

RESULT 10
ID Q9QEW7 PRELIMINARY; PRT; 550 AA.
AC Q9QEW7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

```

```

DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=SEA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AFI19438; AAF02703.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59333 MW; 086E51FED5582BBA CRC64;

Query Match      29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302

RESULT 11
ID Q9WKK4 PRELIMINARY; PRT; 550 AA.
AC Q9WKK4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=WTF;
RC MEDLINE=93329215; PubMed=10400788;
RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
RA Schneider-Schaulies S.;
RT "A recombinant measles vaccine virus expressing wild-type
RT glycoproteins : consequences for viral spread and cell tropism.";
RL J. Virol. 73:6903-6915(1999).
DR EMBL; AJ133108; CAB38075.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59580 MW; 825549996B5D862 CRC64;

Query Match      29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302

```

```

RESULT 12
Q89495 PRELIMINARY; PRT; 550 AA.
AC Q89495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230209; PubMed=1566568;
RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT "Genetic variability of the glycoprotein genes of current wild-type
RT measles isolates.";
RL Virology 188:135-142 (1992).
DR EMBL; M81903; AAA46422.1; -.
DR EMBL; M81901; AAA46421.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59564 MW; A78EC9CD6268E58 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 288 LSEIKGVIVHRLGV 302

RESULT 13
Q8V049 PRELIMINARY; PRT; 550 AA.
AC Q8V049;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G954;
RX MEDLINE=21635526; PubMed=11773423;
RA Waku Koumou D., Wild T.F.;
RT "Adaptation of wild-type measles virus to tissue culture.";
RL J. Virol. 76:1505-1509 (2002).
DR EMBL; AY059392; RAL29688.1; -.
DR PIR; PQ0376; PQ0376.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59551 MW; 9A7A4BA99B4DA8E9 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 288 LSEIKGVIVHRLGV 302

```

```

Db 288 LSEIKGVIVHRLGV 302

RESULT 14
Q9YJ94 PRELIMINARY; PRT; 550 AA.
AC Q9YJ94;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=9301V;
RX MEDLINE=98440529; PubMed=9765410;
RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA Asakawa M., Nagai Y.;
RT "Measles virus attenuation associated with transcriptional impediment
RT and a few amino acid changes in the polymerase and accessory
RT proteins.";
RL J. Virol. 72:8690-8696 (1998).
DR EMBL; AB012949; BAA33877.1; -.
DR EMBL; AB012948; BAA33871.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59512 MW; 7AA4FDD11797BF9 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 288 LSEIKGVIVHRLGV 302

RESULT 15
Q9QEX1 PRELIMINARY; PRT; 550 AA.
AC Q9QEX1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Masusako;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: Deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179430; AAF02695.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.

```


SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;
Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
| | | | | | | | | | | | | | | | | | | | | |
Db 288 LSEIKGVIVHRLEGV 302

Search completed: March 10, 2004, 09:25:38
Job time : 46.4514 secs

Result No.	Score	Query		length	DB	ID	Description
		Match	%				
1	246	99.2	47	5	AAU11428	AAU11428	Synthetic
2	173	69.8	34	5	AAU11424	AAU11424	Synthetic
3	159	64.1	31	5	AAU11420	AAU11420	Synthetic
4	147.5	59.5	46	5	AAU11430	AAU11430	Synthetic
5	144	58.1	51	5	AAU11431	AAU11431	Synthetic
6	141.5	57.1	50	5	AAU11429	AAU11429	Synthetic
7	117	47.2	25	2	AAAG2705	AAAG2705	LHRH-cont
8	117	47.2	42	2	AAAG2708	AAAG2708	LHRH-cont
9	116	46.8	27	2	AAAG2707	AAAG2707	LHRH-cont
10	115	46.8	27	3	AAV68567	AAV68567	Peptide i
11	115	46.8	27	3	AAV91156	AAV91156	MVP Th ep
12	115	46.8	45	2	AAAG2721	AAAG2721	LHRH-cont
13	116	46.8	45	7	ADD89949	ADD89949	LHRH pept
14	113	45.6	47	3	AAV91163	AAV91163	Modified
15	110	44.4	31	3	AAV91175	AAV91175	Modified
16	108	43.5	27	3	AAV91161	AAV91161	Modified
17	108	43.5	27	3	AAV91167	AAV91167	Modified
18	108	43.5	45	3	AAV68573	AAV68573	Peptide i
19	108	43.5	45	3	AAV91165	AAV91165	Modified
20	107	43.1	31	3	AAV91179	AAV91179	Modified
21	106	42.7	28	3	AAV91158	AAV91158	Modified
22	106	42.7	40	2	AAAG7591	AAAG7591	Synthetic
23	106	42.7	40	3	AAV79986	AAV79986	Measles v
24	106	42.7	40	6	ADA25172	ADA25172	Chimeric
25	106	42.7	40	7	ADC89661	ADC89661	H. influe

PA (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.
 XX Claim 11; Page 11; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX Sequence 47 AA;
 SQ Query Match 99.2%; Score 246; DB 5; Length 47;
 Best Local Similarity 100.0%; Pred. No. 4.7e-25;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLGEGVPSLHWSYGLRP 46
 DB 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLGEGVPSLHWSYGLRP 46
 RESULT 2
 AAU11424
 ID AAU11424 standard; peptide; 34 AA.
 XX AAU11424;
 XX 12-MAR-2002 (first entry)
 XX Synthetic immunogen peptide 5.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Measles virus.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 FH Peptide 1. .10
 FT /note= "Gonadotropin releasing hormone epitope"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT Peptide 11. .16
 FT /note= "Spacer peptide"
 FT Peptide 17. .34
 FT /note= "Measles virus fusion protein F epitope"
 XX WO200185763-A2.
 XX 15-NOV-2001.
 XX

PF 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.
 XX Claim 11; Page 9; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX Sequence 34 AA;
 SQ Query Match 69.8%; Score 173; DB 5; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLGEGV 34
 DB 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLGEGV 34
 RESULT 3
 AAU11420
 ID AAU11420 standard; peptide; 31 AA.
 XX AAU11420;
 XX 12-MAR-2002 (first entry)
 XX Synthetic immunogen peptide 1.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Measles virus.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 FH Peptide 1. .18 "Measles virus fusion protein F epitope"
 FT /note= "Spacer peptide"
 FT Peptide 19. .22
 FT /note= "Spacer peptide"
 FT Peptide 23. .31
 FT /note= "Gonadotropin releasing hormone epitope"
 FT Modified-site 31
 FT /note= "Amidated glycine or glycineamide"
 XX WO200185763-A2.

```

XX PD 15-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014363.
XX PR 05-MAY-2000; 2000US-0202328P.
XX PA (APHT-) APHTON CORP.
XX PI Grimes S, Michaeli D, Stevens VC;
XX PN WPI; 2002-049440/06.
XX PD Novel synthetic immunogen for inducing immune response against
XX PT Gonadotropin releasing hormone, comprises fusion peptide having
XX PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX PT or its analog.
XX PS Claim 11; Page 7; 43pp; English.
XX CC The invention relates to a synthetic immunogen for inducing specific
XX CC antibodies against gonadotropin releasing hormone (GnRH) also known as
XX CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
XX CC which comprises a promiscuous helper T-cell peptide epitope and
XX CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX CC useful inducing an immune response against GnRH in an animal subject, and
XX CC as such is useful as a contraceptive and in the treatment of diseases
XX CC such as cancer (of the breast, uterus and other gynaecological cancer),
XX CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX CC prostate cancer. The immunogen is effective in eliciting high and
XX CC specific anti-GnRH antibody titres. The present sequence is a synthetic
XX CC immunogen of the invention
XX SQ Sequence 31 AA;
Query Match 64.1%; Score 159; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.3e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 17 KLLSEIKGVIVHRLEGVSGPSLHWSYGLRP 46
Db 1 KLLSEIKGVIVHRLEGVSGPSLHWSYGLRP 30
RESULT 4
AAU11430
ID AAU11430 standard; peptide; 46 AA.
XX AC AAU11430;
XX DT 12-MAR-2002 (first entry)
XX XX Synthetic immunogen peptide 11.
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX OS Clostridium tetani.
XX OS Mammalia.
XX OS Synthetic.
XX OS Chimeric.
XX Key Location/Qualifiers
XX Peptide 1..10
XX /note= "Gonadotropin releasing hormone epitope (1..10
XX aa)"
XX Misc-difference 1
XX /label= OTHER
XX /note= "Other= Pyro-glutamic acid or 5-oxo proline"
XX Peptide 11..16

```

```

FT Peptide /note= "Spacer peptide"
FT 17..31
FT /note= "Tetanus toxoid (830-844 aa)"
FT Peptide 32..37
FT /note= "Spacer peptide"
FT Peptide 38..46
FT /note= "Gonadotropin releasing hormone epitope (2-10
FT aa)"
FT Modified-site 46
FT /note= "Amidated glycine or glycineamide"
XX PN WO200185763-A2.
XX 15-NOV-2001.
XX 04-MAY-2001; 2001WO-US014363.
XX 05-MAY-2000; 2000US-0202328P.
XX (APHT-) APHTON CORP.
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX PS Claim 11; Page 12; 43pp; English.
XX CC The invention relates to a synthetic immunogen for inducing specific
XX CC antibodies against gonadotropin releasing hormone (GnRH) also known as
XX CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
XX CC which comprises a promiscuous helper T-cell peptide epitope and
XX CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX CC useful inducing an immune response against GnRH in an animal subject, and
XX CC as such is useful as a contraceptive and in the treatment of diseases
XX CC such as cancer (of the breast, uterus and other gynaecological cancer),
XX CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX CC prostate cancer. The immunogen is effective in eliciting high and
XX CC specific anti-GnRH antibody titres. The present sequence is a synthetic
XX CC immunogen of the invention
XX SQ Sequence 46 AA;
Query Match 59.5%; Score 147.5; DB 5; Length 46;
Best Local Similarity 62.5%; Pred. No. 5e-12;
Matches 30; Conservative 4; Mismatches 7; Indels 7; Gaps 2;
Oy 2 HWSYGLRPGSGPSLKL---SEIKGVIVHRLEGVSGPSLHWSYGLRP 46
Db 2 HWSYGLRPGSGPSLQYKANSKFTGI---TELSSGPSLHWSYGLRP 45
RESULT 5
AAU11431
ID AAU11431 standard; peptide; 51 AA.
XX AC AAU11431;
XX 12-MAR-2002 (first entry)
XX Synthetic immunogen peptide 12.
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX OS Plasmodium falciparum.

```

OS Mammalia.
OS Synthetic.
XX Chimeric.

PH Key
XX Location/Qualifiers
FT 1..10
FT /note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
FT
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT 11..16
FT /note= "Spacer peptide"
FT
FT 17..36
FT /note= "Malaria CSP protein (378-398 aa)"
FT 37..42
FT /note= "Spacer peptide"
FT 43..51
FT /note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
FT
FT Modified-site 51
FT /note= "Amidated glycine or glycinamide"
FT
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against gonadotropin releasing hormone, comprises fusion peptide having promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analog.
XX
XX Claim 11; Page 12-13; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific antibodies against gonadotropin releasing hormone (GnRH) also known as luteinising hormone releasing hormone, LH/RH) comprising a fusion peptide which comprises a promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analogue. The synthetic immunogen is useful inducing an immune response against GnRH in an animal subject, and as such is useful as a contraceptive and in the treatment of diseases such as cancer (of the breast, uterus and other gynaecological cancer), endometriosis, uterine fibroids, benign prostatic hypertrophy and prostate cancer. The immunogen is effective in eliciting high and specific anti-GnRH antibody titres. The present sequence is a synthetic immunogen of the invention

Sequence 51 AA;
Query Match 58.1%; Score 144; DB 5; Length 51;
Best Local Similarity 57.1%; Pred. No. 1.7e-11;
Matches 28; Conservative 8; Mismatches 9; Indels 4; Gaps 2;

OY 2 HWSYGLRFGSGPSL--KLISEIK--GVIVHRLGVEGSPSLHWSYGLRP 46
|||||
DB 2 HWSYGLRFGSGPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
|||||

RESULT 6
AAU11429
ID AAU11429 standard; peptide; 50 AA.
XX
AC AAU11429;

XX
DT
XX
DE
XX
XX Synthetic immunogen peptide 10.
XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen; luteinising hormone releasing hormone; LH/RH; contraceptive; promiscuous helper T-cell peptide epitope; immunomimic peptide; breast cancer; uterine cancer; gynaecological cancer; endometriosis; uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX Clostridium tetani.
OS Mammalia.
OS Synthetic.
OS Chimeric.

XX Key
XX Location/Qualifiers
FT 1..10
FT /note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
FT
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT 11..16
FT /note= "Spacer peptide"
FT 17..37
FT /note= "Tetanus toxoid (947-967 aa)"
FT 38..41
FT /note= "Spacer peptide"
FT 42..50
FT /note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
FT
FT Modified-site 50
FT /note= "Amidated glycine or glycinamide"
FT
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against gonadotropin releasing hormone, comprises fusion peptide having promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analog.
XX
XX Claim 11; Page 11; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific antibodies against gonadotropin releasing hormone (GnRH) also known as luteinising hormone releasing hormone, LH/RH) comprising a fusion peptide which comprises a promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analogue. The synthetic immunogen is useful inducing an immune response against GnRH in an animal subject, and as such is useful as a contraceptive and in the treatment of diseases such as cancer (of the breast, uterus and other gynaecological cancer), endometriosis, uterine fibroids, benign prostatic hypertrophy and prostate cancer. The immunogen is effective in eliciting high and specific anti-GnRH antibody titres. The present sequence is a synthetic immunogen of the invention

Sequence 50 AA;
Query Match 57.1%; Score 141.5; DB 5; Length 50;
Best Local Similarity 58.8%; Pred. No. 3.5e-11;

CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
 CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX Sequence 42 AA;

Query Match 47.2%; Score 117; DB 2; Length 42;
 Best Local Similarity 82.8%; Pred. No. 4.9e-08;
 Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLSIKGVIVHRLEGVEGSLHWSYGLRP 46
 :|||||
 Db 15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41

RESULT 9
 AAR62707
 ID AAR62707 standard; peptide; 27 AA.

XX AAR62707;
 AC
 XX 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 XX

DE LHRH-containing immunogenic peptide.

XX Helper T cell epitope; universal immune stimulator; invasin; hapten;
 XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.

XX Synthetic.

XX Key Location/Qualifiers
 FH Domain 1..15
 FT /note= "measles virus F protein helper T cell epitope"
 FT 18..27
 FT /note= "LHRH hapten"
 XX

XX W09425060-A1.

XX 10-NOV-1994.

XX 28-APR-1994; 94WO-US004832.

XX 27-APR-1993; 93US-00057166.

XX 14-APR-1994; 94US-00229275.

XX (LADD//) LADD A E.

XX (WANG//) WANG C Y.

XX (ZAMB//) ZAMB T.

XX Ladd AE, Wang CY, Zamb T;

XX WPI; 1994-357910/44.

XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
 suppress LHRH activity in males and females.

XX Claim 8, 12; Page 86; 213pp; English.

XX Synthetic immunogenic peptides are provided in which a universal immune
 stimulator is linked to a peptide or protein hapten containing B cell
 and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 potent immune responses to the coupled peptide or protein. The stimulator
 consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 immune response to the coupled peptide in members of a heterogeneous
 population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 sequence from the invasin protein of Yersinia. Spacer amino acid
 sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 domains and between the immune stimulator and hapten components. When the

CC hapten is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing, invasin-free immunogenic peptide as above which can be used
 CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
 CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
 CC cancer, or for induction of infertility. This sequence is particularly
 CC preferred. (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 27 AA;

Query Match 46.8%; Score 116; DB 2; Length 27;
 Best Local Similarity 85.7%; Pred. No. 3.9e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVEGSLHWSYGLRP 46
 :|||||
 Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 10

AAAY68567

ID AAY68567 standard; peptide; 27 AA.

XX AAY68567;

XX 05-MAY-2000 (first entry)

DE Peptide immunogen comprising a Th epitope and LHRH target antigen.

XX Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; chimera;
 KW immunocastration.

XX Measles virus.

OS Unidentified.

OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..15 /note= "helper Th epitope AAY68540"

FT Peptide 16..17 /note= "spacer"

FT Peptide 18..27 /note= "LHRH antigenic epitope AAY68566"

XX W09966952-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013960.

XX 20-JUN-1998; 98US-00100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinizing hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 treatment of cancer.

XX Example 1; Page 63; 102pp; English.

XX The present sequence represents a peptide immunogen comprising a helper T
 CC cell (Th) epitope of the F protein of the Measles virus and a target
 CC antigen, luteinising hormone-releasing hormone (LHRH). The peptide

CC immunogens cause induction of a specific immune response to LHRH which is
 CC involved in regulation of spermatogenesis, ovulation, oestrus, sexual
 CC development and secretion of sex hormones. Provision of a promiscuous T
 CC helper epitope (which is functional in genetically diverse subjects)
 CC provides optimum immunogenicity to the B cell epitopes of the target
 CC antigen and thus high antibody titres against the target antigen. The
 CC peptide immunogens of the invention are used to vaccinate against
 CC mammalian LHRH, for use as (reversible) contraceptive; control of hormone
 CC -dependent tumours (cancer of prostate or breast, also endometriosis); to
 CC prevent boar taint (and improve meat quality) and for immunocastration
 CC
 XX Sequence 27 AA;

Query Match 46.8%; Score 116; DB 3; Length 27;
 Best Local Similarity 85.7%; Pred. No. 3.9e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVGPGSLHWSYGLRP 46
 DB 1 LSEIKGVIVHRLGVGPGSLHWSYGLRP 26

RESULT 11
 ID AAY91156
 XX AAY91156 standard; peptide; 27 AA.

AC AAY91156;

XX 12-SEP-2003 (revised)
 DT 22-MAY-2000 (first entry)

DE MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Measles virus.
 OS Rattus sp.
 OS Chimeric.

PN WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013975.

XX 20-JUN-1998; 98US-00100412.

PA (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

PI WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus.

XX Example 1; Page 77; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of

CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and immunocastration)
 CC ; for promoting the growth of animals; or for treating allergies or
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
 CC genetically diverse subjects) into an immunogen improves capacity to
 CC induce a strong T helper cell-mediated immune response, resulting in
 CC production of antibodies against a target antigen. Th can replace carrier
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
 CC represents a promiscuous T helper epitope from the measles virus F (MVF)
 CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
 CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
 CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
 CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
 CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
 CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
 CC epitope. Somatostatin immunogens may be used to promote growth in
 CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
 CC AAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may
 CC be used to prevent HIV infection of T cells. AAY90212 is a modified
 CC version of a human IGE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
 CC are Th epitope/IGE CH3 antigenic peptides which may be used in the
 CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
 CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
 CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
 CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
 CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
 CC AAY91228-Y91231 represent CPTP-derived peptides and AAY91232-Y91241 are
 CC immunogens comprising a CPTP peptide and a Th epitope which may be used
 CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
 CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
 CC Y91251 and AAY91253-Y91273 are antigenic peptides comprising MVF Th and
 CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
 CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
 CC invasion protein epitope from *Yersinia* species, and hinge spacer peptide,
 CC both of which may optionally be used in the antigenic peptides of the
 CC invention. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 27 AA;

Query Match 46.8%; Score 116; DB 3; Length 27;
 Best Local Similarity 85.7%; Pred. No. 3.9e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVGPGSLHWSYGLRP 46
 DB 1 LSEIKGVIVHRLGVGPGSLHWSYGLRP 26

RESULT 12
 AAR62721

ID AAR62721 standard; peptide; 45 AA.

XX AAR62721;

XX 25-MAR-2003 (revised)

DT 10-SEP-1995 (first entry)

XX LHRH-containing immunogenic peptide.

XX Helper T cell epitope; universal immune stimulator; invasin; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.

XX Synthetic.

XX Key Location/Qualifiers
 FH Domain 1. .16
 FT

PT FT Domain /note= "invasin domain"
 FT 19...33
 FT /note= "measles virus F protein helper T cell epitope"
 FT 36...45
 FT /note= "LHRH hapten"
 XX WO9425060-A1.
 PN 10-NOV-1994.
 PD
 XX 28-APR-1994; 94WO-US004832.
 XX 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.
 XX (LADD/) LADD A. E.
 PA (WANG/) WANG C. Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 PI WPI; 1994-357910/44.
 DR
 XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 PT
 XX Claim 8; Page 88; 213pp; English.
 PS
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasins protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasins and Th
 CC domains and between the immune stimulator and hapten components. When the
 CC hapten is LHRH, then optionally the invasins domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing immunogenic peptide as above which can be used as a potent
 CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
 CC carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis,
 CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX SQ Sequence 45 AA;
 Query Match 46.8%; Score 116; DB 2; Length 45;
 Best Local Similarity 85.7%; Pred. No. 7.3e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 19 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 46
 DB 19 LSEIKGVIVHRLGVEGSGE--HWSYGLRP 44
 RESULT 13
 ADD89949
 ID ADD89949 standard; protein; 45 AA.
 XX
 AC ADD89949;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX LHRH peptide used in immunostimulant complex for prostate cancer vaccine.
 DE
 XX Immunostimulant; vaccine; human; immunogen; LHRH; immunotherapy;
 KW prostate cancer.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX WO2003068169-A2.
 PN 21-AUG-2003.
 PD
 XX 14-FEB-2003; 2003WO-US004711.
 PF 14-FEB-2002; 2002US-00076674.
 PR 31-JAN-2003; 2003US-00076674.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA
 XX Sokoll KK;
 XX WPI; 2003-778890/73.
 DR
 XX Stabilized immunostimulating complex, useful for vaccination, e.g.
 PT against human immune deficiency viruses, comprises cationic peptide
 PT immunogen and anionic oligonucleotide.
 XX
 XX Claim 17; SEQ ID NO 9; 159pp; English.
 PS
 XX The present sequence is that of a synthetic immunogenic peptide derived
 CC from human LHRH. This is an example of peptides that can be used in
 CC claimed immunostimulatory complexes of the invention that are
 CC specifically adapted to act as adjuvant and as peptide immunogen
 CC stabiliser. The complexes comprise a Cpg oligonucleotide and a
 CC biologically active peptide immunogen. The complex is particulate and can
 CC efficiently present peptide immunogens to the cells of the immune system
 CC to produce an immune response. The complexes may be prepared with various
 CC ratios of peptides to Cpg oligonucleotides to provide different physical
 CC properties, such as the size of the microparticle. An immunostimulatory
 CC complex comprising the present LHRH derived peptide can be used in a
 CC vaccine for prostate cancer.
 XX
 XX SQ Sequence 45 AA;
 Query Match 46.8%; Score 116; DB 7; Length 45;
 Best Local Similarity 85.7%; Pred. No. 7.3e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 19 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 46
 DB 19 LSEIKGVIVHRLGVEGSGE--HWSYGLRP 44
 RESULT 14
 AAY91163
 ID AAY91163 standard; peptide; 27 AA.
 XX
 AC AAY91163;
 XX
 DT 12-SEP-2003 (revised)
 DT 22-MAY-2000 (first entry)
 XX
 DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.
 XX
 KW Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; anti-malarial; CEMP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.
 XX
 OS Measles virus.
 OS Rattus sp.
 OS Chimeric.
 XX
 XX WO9966957-A2.
 PN
 XX 29-DEC-1999.
 PD
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 36.9416 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248
Sequence: 1 XHSYGLRPGSSGSLKLS.....HRLEGVEGSLHWSYGLRFX 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	246	99.2	47	9	US-09-848-834A-17
2	173	69.8	34	9	US-09-848-834A-13
3	159	64.1	31	9	US-09-848-834A-9
4	147.5	59.5	46	9	US-09-848-834A-19
5	144	58.1	51	9	US-09-848-834A-20
6	141.5	57.1	50	9	US-09-848-834A-18
7	116	46.8	45	14	US-10-076-674-9
8	116	46.8	45	15	US-10-355-161A-9
9	106	42.7	40	14	US-10-223-711-10
10	102	41.1	75	10	US-09-847-102A-31
11	102	41.1	75	14	US-10-285-976-229
12	101	40.7	75	10	US-09-847-102A-33
13	101	40.7	75	14	US-10-285-976-231
14	96.5	38.9	49	9	US-09-019-010-4
15	96.5	38.9	49	10	US-09-305-924-11

16	96.5	38.9	695	10	US-09-305-924-13	Sequence 13, Appl
17	96	38.7	40	14	US-10-223-711-11	Sequence 11, Appl
18	94.5	38.1	40	10	US-09-964-201A-35	Sequence 35, Appl
19	94.5	38.1	41	10	US-09-964-201A-34	Sequence 34, Appl
20	89	35.9	31	9	US-09-848-834A-15	Sequence 15, Appl
21	87	35.1	36	9	US-09-848-834A-16	Sequence 16, Appl
22	87	35.1	37	9	US-09-848-834A-14	Sequence 14, Appl
23	86.5	34.9	20	10	US-09-964-201A-26	Sequence 26, Appl
24	86.5	34.9	20	10	US-09-964-201A-29	Sequence 29, Appl
25	86.5	34.9	20	10	US-09-964-201A-30	Sequence 30, Appl
26	86.5	34.9	20	10	US-09-964-201A-31	Sequence 31, Appl
27	79	31.9	34	9	US-09-848-834A-10	Sequence 10, Appl
28	77	31.0	16	9	US-09-848-834A-8	Sequence 8, Appl
29	77	31.0	25	15	US-10-411-544-32	Sequence 32, Appl
30	73	29.4	28	9	US-09-848-834A-11	Sequence 11, Appl
31	73	29.4	33	9	US-09-848-834A-12	Sequence 12, Appl
32	72	29.0	15	10	US-09-747-802-16	Sequence 16, Appl
33	72	29.0	15	10	US-09-747-802-30	Sequence 30, Appl
34	72	29.0	15	10	US-09-865-294-8	Sequence 8, Appl
35	72	29.0	15	10	US-09-865-294-22	Sequence 22, Appl
36	72	29.0	15	14	US-10-281-446-20	Sequence 20, Appl
37	72	29.0	15	15	US-10-411-544-10	Sequence 10, Appl
38	72	29.0	36	14	US-10-351-641-505	Sequence 505, App
39	72	29.0	550	9	US-09-873-233A-18	Sequence 18, Appl
40	72	29.0	550	9	US-09-873-233A-20	Sequence 20, Appl
41	71	28.6	17	10	US-09-305-924-4	Sequence 4, Appl
42	69	27.8	15	10	US-09-747-802-37	Sequence 37, Appl
43	69	27.8	15	10	US-09-865-294-29	Sequence 29, Appl
44	69	27.8	16	14	US-10-223-711-3	Sequence 3, Appl
45	69	27.8	19	10	US-09-747-802-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; TYPE: PRT
; LENGTH: 47
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid se
; OTHER INFORMATION: uence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (47)..(47)
; OTHER INFORMATION: Amidated-glycine or glycynamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein

```

; NAME/KEY: PEPTIDE
; LOCATION: (35)..(38)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (39)..(47)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match          99.2%; Score 246; DB 9; Length 47;
Best Local Similarity 100.0%; Pred. No. 4.3e-24;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 HWSYGLRPGSSGPSLKLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46
    |||||
Db   2 HWSYGLRPGSSGPSLKLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46

RESULT 2
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of b
; OTHER INFORMATION: he Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
; OTHER INFORMATION: virus fusion protein, F
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

Query Match          69.8%; Score 173; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.7e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 HWSYGLRPGSSGPSLKLSEIKGVIVHRLEGVE 34
    |||||
Db   2 HWSYGLRPGSSGPSLKLSEIKGVIVHRLEGVE 34

RESULT 3
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the M
; OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to amin
; OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated Lysine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
; OTHER INFORMATION: measles virus fusion protein, F
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(22)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (23)..(31)
; OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
; OTHER INFORMATION: GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (31)..(31)
; OTHER INFORMATION: Amidated glycine or glycineamide
US-09-848-834A-9

Query Match          64.1%; Score 159; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.1e-13;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46
    |||||
Db   1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30

RESULT 4
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanu
; OTHER INFORMATION: toxoid precursor (Tetoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE

```

```

; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

Query Match          59.5%; Score 147.5; DB 9; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.4e-11;
Matches 30; Conservative 4; Mismatches 7; Indels 7; Gaps 2;

QY 2 HWSYGLRPGSGPSLKL--SEIKGVHRLGVGEGPSLHWSYGLRP 46
DB 2 HWSYGLRPGSGPSLQVYKANSKFIGI-----TELSGPGSLHWSYGLRP 45

RESULT 5
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immuncogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmod
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          58.1%; Score 144; DB 9; Length 51;
Best Local Similarity 57.1%; Pred. No. 4.5e-11;
Matches 28; Conservative 8; Mismatches 9; Indels 4; Gaps 2;

```

```

QY 2 HWSYGLRPGSGPSL--KLLSEIK--GVIVHRLGVGEGPSLHWSYGLRP 46
DB 2 HWSYGLRPGSGPSLDEKXIARKKASVFNWVNSSGPGSLHWSYGLRP 50

RESULT 6
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immuncogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the T
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (50)..(50)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Te
; OTHER INFORMATION: oxylysine
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

Query Match          57.1%; Score 141.5; DB 9; Length 50;
Best Local Similarity 58.8%; Pred. No. 9.1e-11;
Matches 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 2 HWSYGLRPGSGPSL-----KLLSEIKGVIVHRLGVGEGPSLHWSYGLRP 46
DB 2 HWSYGLRPGSGPSLNNFTVFWLRPKVSASHL---EGPSLHWSYGLRP 49

RESULT 7
US-10-076-674-9
; Sequence 9, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1

```

Qy	6	GLRPGSSGSLKLSLSEIKGVIVHRLGEV	34
Db	12	GTRDHKKGPSLKLSLIKGVIHRLGEV	40
US-10-076-674-9			
Query Match 46.8%; Score 116; DB 14; Length 45;			
Best Local Similarity 85.7%; Pred. No. 1.4e-07;			
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;			
Qy	19	LSEIKGVIVHRLGEGVSGSLHSYGLRP	46
Db	19	LSEIKGVIVHRLGEGVGGE--HWSYGLRP	44
US-10-355-161A-9			
RESULT 8			
US-10-355-161A-9			
Sequence 9, Application US/10355161A			
Publication No. US2004000987A1			
GENERAL INFORMATION:			
APPLICANT: Sokoll, Kenneth K.			
TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System			
FILE REFERENCE: Immunogen Delivery System			
CURRENT APPLICATION NUMBER: US/10/355,161A			
CURRENT FILING DATE: 2003-01-31			
PRIOR APPLICATION NUMBER: US 10/076674			
PRIOR FILING DATE: 2002-02-14			
NUMBER OF SEQ ID NOS: 13			
SOFTWARE: Patent in version 3.1			
SEQ ID NO 9			
LENGTH: 45			
TYPE: PRT			
ORGANISM: Human			
OTHER INFORMATION: synthetic construct			
US-10-355-161A-9			
Query Match 46.8%; Score 116; DB 15; Length 45;			
Best Local Similarity 85.7%; Pred. No. 1.4e-07;			
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;			
Qy	19	LSEIKGVIVHRLGEGVSGSLHSYGLRP	46
Db	19	LSEIKGVIVHRLGEGVGGE--HWSYGLRP	44
US-10-223-711-10			
RESULT 9			
US-10-223-711-10			
Sequence 10, Application US/10223711			
Publication No. US2003011334A1			
GENERAL INFORMATION:			
APPLICANT: Bakaletz, Lauren O.			
TITLE OF INVENTION: Synthetic Chimeric Fibroin Peptides			
FILE REFERENCE: 18525/04058			
CURRENT APPLICATION NUMBER: US/10/223,711			
CURRENT FILING DATE: 2002-08-19			
PRIOR APPLICATION NUMBER: 09/148,711			
PRIOR FILING DATE: 1998-09-04			
PRIOR APPLICATION NUMBER: 08/460,502			
PRIOR FILING DATE: 1995-06-02			
NUMBER OF SEQ ID NOS: 12			
SOFTWARE: Patent in version 3.1			
SEQ ID NO 10			
LENGTH: 40			
TYPE: PRT			
ORGANISM: Artificial			
FEATURE:			
OTHER INFORMATION: synthetic construct			
US-10-223-711-10			
Query Match 42.7%; Score 106; DB 14; Length 40;			
Best Local Similarity 79.3%; Pred. No. 2.3e-06;			
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;			

```
Query Match          41.1%; Score 102; DB 14; Length 75;
Best Local Similarity 47.2%; Pred. No. 1.5e-05;
Matches 25; Conservative 2; Mismatches 2; Indels 24; Gaps 1;

QY 6 GLRPGSS-----GPSLKLLSEIKGVIVHRLEGVE 34
   |||||
Db 23 GLQFAGGTGGGGGAPPYATLEHPFHCGPSKLLSLIKGVIVHRLEGVE 75

RESULT 12
US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Carr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leoni M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
; FILE REFERENCE: 22000-20629.00
; CURRENT APPLICATION NUMBER: US/09/847,102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PMMVF-FZD2
US-09-847-102A-33

Query Match          40.7%; Score 101; DB 10; Length 75;
Best Local Similarity 91.3%; Pred. No. 2.1e-05;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIVHRLEGVEGPSL 38
   |||||
Db 1 MKLLSLIKGVIVHRLEGVEGPSL 23

RESULT 13
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Carr, Maripat
; APPLICANT: Carson, Dennis A.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; FILE REFERENCE: 023070-130320US
; CURRENT APPLICATION NUMBER: US/10/285,976
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PMMVF-ZD2
```

```
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-231

Query Match          40.7%; Score 101; DB 14; Length 75;
Best Local Similarity 91.3%; Pred. No. 2.1e-05;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIVHRLEGVEGPSL 38
   |||||
Db 1 MKLLSLIKGVIVHRLEGVEGPSL 23

RESULT 14
US-09-019-010-4
; Sequence 4, Application US/09019010
; Patent No. US20010014330A1
; GENERAL INFORMATION:
; APPLICANT: HARLAND, RICHARD
; APPLICANT: MANNS, JOHN G.
; APPLICANT: ACRES, STEPHEN D.
; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
; MOLECULES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,010
; FILING DATE: 05-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,883
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-010-4

Query Match          38.9%; Score 96.5; DB 9; Length 49;
Best Local Similarity 44.7%; Pred. No. 4.7e-05;
Matches 21; Conservative 1; Mismatches 10; Indels 15; Gaps 2;

QY 2 HWSYGLRPGSSGFSLKLLSEIKGVIVHRLEGVE--GPSLHWSYGLRP 46
   |||||
Db 2 HWSYGLRPGSGSQDWSY-----GLRPGSGSQHWSYGLRP 35

RESULT 15
US-09-305-924-11
; Sequence 11, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
```


APPLICANT: Stephen D. Acres
APPLICANT: Richard Harland
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,217
EARLIER FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 11
LENGTH: 49
TYPE: PRT
ORGANISM: GRNH
US-09-305-924-11

Query Match 38.9%; Score 96.5; DB 10; Length 49;
Best Local Similarity 44.7%; Pred. No. 4.7e-05;
Matches 21; Conservative 1; Mismatches 10; Indels 15; Gaps 2;
QY 2 HWSYGLRPGSSGSLKLLSEIKGVIVHRLGVE--GPSLHWSYGLRP 46
|||||
Db 2 HWSYGLRPGSSGSDWSY-----GLRPGSSQHWSYGLRP 35
|||||

Search completed: March 10, 2004, 10:25:49
Job time : 36.9416 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 13.9377 seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSSGSLDEKKIAMKXASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB pep.*
- 2: /cgn2_6/ptodata/2/iaa/5E COMB pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	49.7	412	1	US-08-313-288B-18
2	91	49.7	423	2	US-08-760-797A-1
3	91	49.7	424	2	US-08-760-797A-3
4	91	49.7	424	3	US-08-932-929B-1
5	91	49.7	424	3	US-08-932-929B-3
6	89	48.6	21	1	US-08-186-266-6
7	89	48.6	21	1	US-08-446-592-48
8	89	48.6	21	2	US-08-488-351A-48
9	89	48.6	21	3	US-09-100-409A-54
10	89	48.6	21	4	US-08-464-496-17
11	89	48.6	21	4	US-08-788-822A-12
12	89	48.6	21	4	US-08-197-484-97
13	89	48.6	21	4	US-09-543-608A-39
14	89	48.6	21	5	PCT-US95-02121-97
15	89	48.6	21	5	PCT-US95-13841-20
16	89	48.6	33	1	US-08-446-692-27
17	89	48.6	33	2	US-08-488-351A-27
18	79	43.2	17	4	US-08-464-496-16
19	79	43.2	17	4	US-08-197-484-96
20	79	43.2	17	5	PCT-US95-02121-96
21	75	41.0	16	2	US-08-817-933A-7
22	73.5	40.2	20	1	US-08-465-167A-20
23	73.5	40.2	20	4	US-08-627-820-20
24	73.5	40.2	20	5	PCT-US92-07218-17
25	71	38.8	17	1	US-08-188-223-6
26	71	38.8	17	3	US-08-968-466-6
27	71	38.8	17	4	US-08-478-546B-6

Patent No. 5169933
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 19, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 18, Appl
Sequence 28, Appl
Sequence 12, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 10, Appl

70 38.3 15 6 5169933-30
699 2 US-08-694-865-16
699 3 US-09-124-491-16
699 4 US-09-383-912-16
699 1 US-08-465-167A-19
63.5 34.7 16 4 US-08-627-820-19
63.5 34.7 16 5 PCT-US92-07218-16
62 33.9 14 1 US-08-453-588-23
62 33.9 14 3 US-08-521-079-23
62 33.9 17 1 US-07-690-983D-18
62 33.9 18 1 US-07-630-983D-28
62 33.9 21 1 US-08-305-871A-12
62 33.9 49 1 US-08-387-156-4
62 33.9 49 2 US-08-694-865-4
62 33.9 49 3 US-08-878-748-4
62 33.9 49 3 US-09-124-491-4
62 33.9 49 4 US-09-383-912-4
62 33.9 544 1 US-08-387-156-10

ALIGNMENTS

RESULT 1
US-08-313-288B-18
; Sequence 18, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Aviuh Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, P-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John F.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-288B-18

Query Match 49.7%; Score 91; DB 1; Length 412;

Best Local Similarity 55.3%; Pred. No. 6.1e-05;

Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LAPGSGPSLD-----EKKIAMKXASSVFNVNS 36

Db 361 IKFGSANKPKDELVDYNDIEKKICKVKCSSVFNVNS 398

```

RESULT 2
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-1
Query Match 49.7%; Score 91; DB 2; Length 423;
Best Local Similarity 55.3%; Pred. No. 6.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRGSSGSPSLD-----EKKIARKMEKASSVFNVNS 36
Db 155 IKPGSANKPKDELVDYNDIEKKICKMEKCSSVFNVNS 192

RESULT 3
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-3
Query Match 49.7%; Score 91; DB 2; Length 424;
Best Local Similarity 55.3%; Pred. No. 6.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRGSSGSPSLD-----EKKIARKMEKASSVFNVNS 36
Db 152 IKPGSANKPKDELVDYNDIEKKICKMEKCSSVFNVNS 189

RESULT 4
US-08-932-929B-1
; Sequence 1, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096

```

```

; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-932-929B-1

Query Match      49.7%; Score 91; DB 3; Length 424;
Best Local Similarity 55.3%; Pred. No. 6.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNVS 36
Db 156 IKPGSANKPKDLDYNDIEKKICKMEKCSVFNVNVS 193

RESULT 5
US-08-932-929B-3
; Sequence 3, Application US/09932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1PWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-932-929B-3

Query Match      49.7%; Score 91; DB 3; Length 424;
Best Local Similarity 55.3%; Pred. No. 6.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNVS 36
Db 152 IKPGSANKPKDLDYNDIEKKICKMEKCSVFNVNVS 189

RESULT 6
US-08-186-266-6
; Sequence 6, Application US/08186266
; Patent No. 5662907
; GENERAL INFORMATION:
; APPLICANT: KUBO, Ralph T.
; APPLICANT: GREY, Howard M.
; APPLICANT: SETTE, Alessandro
; APPLICANT: CELIS, Esteban
; TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
; TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
; TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,266
; FILING DATE: 25-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/159,339
; FILING DATE: 29-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-50-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Plasmodium falciparum CS
; OTHER INFORMATION: protein at positions 378-398."
US-08-186-266-6

Query Match      48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKIKAKMEKASSVFNVNVS 36
Db 3 EKIKAKMEKASSVFNVNVS 21

RESULT 7
```

```
US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Chang Yi
; APPLICANT: Wang, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-48

Query Match 48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAXKXKASSVFNVNS 36
Db 3 EKXIAXKXKASSVFNVNS 21

RESULT 8
US-08-488-351A-48
; Sequence 48, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25

US-09-848-834a-16.open.ra1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-48

Query Match 48.6%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAXKXKASSVFNVNS 36
Db 3 EKXIAXKXKASSVFNVNS 21

RESULT 9
US-09-100-409A-54
; Sequence 54, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
```

```
/ TELEFAX: 212-751-6849
/ INFORMATION FOR SEQ ID NO: 54:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 21 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-09-100-409A-54

Query Match 48.6%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNVS 36
DB 3 EKXIAKMEKASSVFNVNVS 21

RESULT 10
US-08-464-496-17
/ Sequence 17, Application US/08464496
/ Patent No. 6322789
/ GENERAL INFORMATION:
/ APPLICANT: Epimmune, Inc.
/ APPLICANT: Vitellio, Maria
/ APPLICANT: Chesnut, Robert
/ TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
/ TITLE OF INVENTION: EPITOPES
/ FILE REFERENCE: 39863-20001.13
/ CURRENT APPLICATION NUMBER: US/08/464,496
/ CURRENT FILING DATE: 1995-06-05
/ PRIOR APPLICATION NUMBER: 07/935,811
/ PRIOR FILING DATE: 1992-08-26
/ PRIOR APPLICATION NUMBER: 07/874,491
/ PRIOR FILING DATE: 1992-04-27
/ PRIOR APPLICATION NUMBER: 07/827,682
/ PRIOR FILING DATE: 1992-01-29
/ PRIOR APPLICATION NUMBER: 07/749,568
/ PRIOR FILING DATE: 1991-08-26
/ NUMBER OF SEQ ID NOS: 75
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Malaria circumsporozoite 378-398
US-08-464-496-17

Query Match 48.6%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNVS 36
DB 3 EKXIAKMEKASSVFNVNVS 21

RESULT 11
US-08-788-822A-12
/ Sequence 12, Application US/08788822A
/ Patent No. 6413935
/ GENERAL INFORMATION:
/ APPLICANT: Alexander, Jeffrey L.
/ APPLICANT: DeFrees, Shawn
/ APPLICANT: Sette, Alessandro
/ TITLE OF INVENTION: Induction of Immune Response Against
/ TITLE OF INVENTION: Desired Determinants
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA

/ TELEFAX: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/788,822A
/ FILING DATE: 23-JAN-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/010,510
/ FILING DATE: 24-JAN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 014137-009210US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 21 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-788-822A-12

Query Match 48.6%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNVS 36
DB 3 EKXIAKMEKASSVFNVNVS 21

RESULT 12
US-08-197-484-97
/ Sequence 97, Application US/08197484
/ Patent No. 6419931
/ GENERAL INFORMATION:
/ APPLICANT: VITIELLO, Maria A.
/ APPLICANT: CHESTNUT, Robert W.
/ APPLICANT: SETTE, Alessandro D.
/ APPLICANT: CELIS, Esteban
/ APPLICANT: GRAY, Howard
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
/ TITLE OF INVENTION: CTL IMMUNITY
/ NUMBER OF SEQUENCES: 153
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend Khourie and Crew
/ STREET: Steuart Street Tower, One Market Plaza
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: US
/ ZIP: 94105-1493
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/197,484
/ FILING DATE: 16-FEB-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/935,811
/ FILING DATE: 26-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/874,491
/ FILING DATE: 27-APR-1992
```

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
US-08-197-484-97

Query Match 48.6%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKXIAKVEKASSVFNVNS 36
Db 3 EKXIAKVEKASSVFNVNS 21

RESULT 13
US-09-543-608A-39
; Sequence 39, Application US/09543608A
; Patent No. 6602510
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Celis, Esteban
; APPLICANT: Keogh, Elissa A.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: HLA Class I A2 Tumor Associated Antigen
; TITLE OF INVENTION: Peptides and Vaccine Compositions
; FILE REFERENCE: 018623-015710US
; CURRENT APPLICATION NUMBER: US/09/543,608A
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 21
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmodium falciparum CS protein positions 378-398
US-09-543-608A-39

Query Match 48.6%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKXIAKVEKASSVFNVNS 36
Db 3 EKXIAKVEKASSVFNVNS 21

RESULT 14
PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match 48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKXIAKVEKASSVFNVNS 36
Db 3 EKXIAKVEKASSVFNVNS 21

RESULT 15
PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Wallfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin

```

, STREET: 345 Park Avenue
, CITY: New York
, STATE: NY
, COUNTRY: USA
, ZIP: 10154
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: Wordperfect 5.1
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: PCT/US95/13841
, FILING DATE: 25-OCT-1995
, CLASSIFICATION:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/328,519
, FILING DATE: 25-OCT-1994
, CLASSIFICATION:
, ATTORNEY/AGENT INFORMATION:
, NAME: Lin, Maria C.H.
, REGISTRATION NUMBER: 29,323
, REFERENCE/DOCKET NUMBER: 1151-4117
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 212-758-4800
, TELEFAX: 212-751-6849
, TELEX: 421792
, INFORMATION FOR SEQ ID NO: 20:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 21 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: peptide
, PCT-US95-13841-20

```

```

Query Match      48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      18 EKXIAKMEKASSVFNVVNS 36
          |||||
Db      3 EKXIAKMEKASSVFNVVNS 21

```

```

Search completed: March 10, 2004, 09:28:56
Job time : 14.9377 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 11.2062 seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSGPSLDEKXIAKMKKASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	53.0	388	2 A39756	circumsporozoite p
2	91	49.7	405	2 S05428	circumsporozoite p
3	91	49.7	412	1 O22QAF	circumsporozoite p
4	91	49.7	424	2 A54533	circumsporozoite p
5	90	49.2	442	2 A54529	circumsporozoite p
6	58	31.7	10	1 RHPGG	gonadoliberin - pl
7	58	31.7	10	1 RHSHG	gonadoliberin - sh
8	58	31.7	67	2 I78541	gonadoliberin prec
9	58	31.7	89	2 I51423	gonadoliberin prec
10	58	31.7	90	1 RHMSG	gonadoliberin prec
11	58	31.7	92	1 RHUG	gonadoliberin prec
12	58	31.7	92	1 RHUG	gonadoliberin prec
13	58	31.7	332	1 O22QMB	circumsporozoite p
14	58	31.7	348	1 O22QBK	circumsporozoite p
15	55	30.1	98	2 I50739	gonadotropin-relea
16	55	30.1	264	2 A44969	circumsporozoite p
17	55	30.1	367	1 O22QMY	circumsporozoite p
18	54	29.5	10	1 RHQOI	gonadoliberin I -
19	54	29.5	92	2 I50644	gonadoliberin I pr
20	53.5	29.2	90	2 A37335	gonadoliberin prec
21	53	29.0	70	2 AC2866	hypothetical prote
22	53	29.0	501	2 T32848	hypothetical prote
23	52.5	28.7	487	2 A59645	methylnalonate-sem
24	52	28.4	719	2 T52510	hypothetical prote
25	52	28.4	1401	2 T48079	hypothetical prote
26	51.5	28.1	80	1 RHID13	gonadoliberin I pr
27	51.5	28.1	90	2 JC7395	salmon-type gonado
28	51.5	28.1	444	1 C64226	trigger factor MG2
29	51	27.9	315	2 F98295	hypothetical prote

conserved hypothet
hypothetical prote
circumsporozoite p
circumsporozoite p
MSH1 protein - yea
gonadoliberin prec
glutathione transf
glutathione trans
succinate-semialde
succinate-semialde
glutathione trans
glutathione trans
probable polygalac
thiamin biosynthes
probable polygalac
thiamin biosynthes
thiamin biosynthes

ALIGNMENTS

RESULT 1

A39756 circumsporozoite protein - Plasmodium reichenowi

C:Species: Plasmodium reichenowi

C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999

C:Accession: A39756

R:Lab, A.A.; Goldman, I.F.

J. Biol. Chem. 266, 6686-6689, 1991

A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malari

A:Reference number: A39756; MUID:91201303; PMID:2016283

A:Accession: A39756

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <LAL>

A:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-366/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match 53.0%; Score 97; DB 2; Length 388;

Best Local Similarity 57.9%; Pred. No. 1.7e-05;

Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPGSAGKPDQDYNDLEKIKCKKSSVFNVNS 36

DB 337 IRPGSAGKPDQDYNDLEKIKCKKSSVFNVNS 374

RESULT 2

S05428 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C:Species: Plasmodium falciparum

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000

C:Accession: S05428; A45527; I60657

R:Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate v

A:Reference number: S05428; MUID:89345189; PMID:2668895

A:Accession: S05428

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <CAM>

A:Cross-references: EMBL:X15363

R:Caspers, P.; Gentz, R.; Maile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate

A:Reference number: A45527; MUID:89364998; PMID:2671723

A:Accession: A45527

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <CAS>

A:Cross-references: GB:M2982; GB:J04650; NID:g160168; PIDN:AAA29527.1; PID:g160169

R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C;Accession: A93780; A01411
B;Burgus R.; Butcher, M.; Amos, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
Proc.Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A;Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor
A;Reference number: A93780; MUID:72094314; PMID:4550508
A;Accession: A93780
A;Molecule type: protein
A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/0/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 31.7%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10
|||||

RESULT 8
I78541
gonadoliberin precursor - rhesus macaque (fragment)
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A;Title: Developmental expression of the genes encoding transforming growth factor alpha
A;Reference number: I58134; MUID:95124501; PMID:7545971
A;Accession: I78541
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-67 <RES>
A;Cross-references: GB:S75918; NID:9912831; PIDN:AA33096.1; PID:g912832
C;Superfamily: gonadoliberin

Query Match 31.7%; Score 58; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 7 HWSYGLRPG 15
|||||

RESULT 9
I51423
gonadoliberin precursor - African clawed frog
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51423
R;Hayes, W.P.; Wray, S.; Battley, J.F.
Endocrinology 134, 1835-1845, 1994
A;Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma
A;Reference number: I51423; MUID:94185563; PMID:8137750
A;Accession: I51423
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:L28040; NID:9496291; PIDN:AAA49728.1; PID:g496292
C;Genetics:
A;Gene: GnRH-I
C;Superfamily: gonadoliberin

Query Match 31.7%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
|||||

RESULT 10
RHMSG
gonadoliberin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasin
N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
A;Reference number: A47578; MUID:8769928; PMID:3024317
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C;Genetics:
A;Insertions: 45/3; 77/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropi
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activit
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;22-31/Product: gonadoliberin #status predicted <GLB>
F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F;22/Modified site: pyroglutamic acid (Gln) (in mature form) #status predict
F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following 9

Query Match 31.7%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 23 HWSYGLRPG 31
|||||

RESULT 11
RHUG
gonadoliberin precursor [validated] - human
N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasin
N;Contains: gonadoliberin-associated protein (GAP); progonadoliberin
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: S05308; A26173; A93342; A90108; A01410; S45718
R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
A;Reference number: S05308; MUID:89366682; PMID:2671939
A;Accession: S05308
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-92 <HAY>
A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: A26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
A;Experimental source: hypothalamus

Cell. Mol. Neurobiol. 12, 447-454, 1992

A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone releasing hormone

A:Reference number: A48410; MUID:93105480; PMID:1468115

A:Accession: A48410

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-92 <MAI>

A:Cross-references: GB:S50870; PIDN:AAB24572.1; PID:G262060

A:Experimental source: thymus

A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIPI:121083)

C:Genetics: 47/3; 79/3

A:Introns: 47/3; 79/3

C:Function:

A:Description: stimulates pituitary secretion of lutropin and follitropin

A:Note: gonadolibirin-associated protein may have prolactin release inhibiting activity

C:Superfamily: gonadolibirin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid;

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-92/Product: progadolibirin #status predicted <PGN>

F:24-33/Product: gonadolibirin #status predicted <GLN>

F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>

F:37/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted <P>

F:23/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gl

Query Match 31.7%; Score 58; DB 1; Length 92;

Best Local Similarity 100.0%; Pred.No. 0.86;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
| | | | | | | | | |

Db 25 HWSYGLRPG 33

RESULT 13

OZZOMB

N:Alternate names: sporozoite surface antigen

C:Species: Plasmodium berghei

C:Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999

C:Accession: A44948; A25083; S13446

R:Lanar, D.E.

Mol. Biochem. Parasitol. 39, 151-154, 1990

A:Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK65

A:Reference number: A44948; MUID:90158693; PMID:2406593

A:Accession: A44948

A:Molecule type: DNA

A:Residues: 1-332 <LAN>

A:Cross-references: GB:M28887

R:Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.

Mol. Cell. Biol. 6, 3965-3972, 1986

A:Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification

A:Reference number: A25083; MUID:87089740; PMID:2432395

A:Accession: A25083

A:Molecule type: DNA

A:Residues: 1-26, 'A', 124-332 <WEB>

A:Cross-references: GB:M14135; NID:G160245; PIDN:AAA29577.1; PID:G160246

R:Weber, J.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hockmeier, J.

Exp. Parasitol. 63, 295-300, 1987

A:Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.

A:Reference number: S13446; MUID:87218962; PMID:3556207

A:Accession: S13446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 61-122, 'A', 124-332 <WEB>

A:Cross-references: GB:M25445; NID:G160177; PIDN:AAA29531.1; PID:G160178

C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

C:Keywords: sporozoite; surface antigen; tandem repeat

F:124-332/Product: signal sequence #status predicted <SIG>

F:94-189/Region: 8-residue repeats

F:193-230/Region: 2-residue repeats

Db 24 HWSYGLSPGGK-RDLDNFSDTLGNMVEEFPRVEAPCSVF 61

Search completed: March 10, 2004, 09:16:51
Job time : 12.265 secs

RESULT 14
OZ03CB
circumsporozoite protein precursor - Plasmodium berghei (strain ANKA clone 2.341)
N;Alternate names: sporozoite surface antigen
C;Species: Plasmodium berghei
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: S07873; S12571
R;Lockyer, M.J.; Davies, C.S.; Suhrbier, A.; Sinden, R.E.
Nucleic Acids Res. 18, 376, 1990
A;Title: Nucleotide sequence of the plasmodium berghei circumsporozoite protein gene fragment
A;Reference number: S07873; MUID:90221834; PMID:2183186

A:Accession: S07873
A:Molecule type: DNA
A:Residues: 1-348 <LOC>
A:Cross-references: EMBL:X17606
R:Lockyer, M.J.
submitted to the EMBL Data Library, November 1989
A:Reference number: S12571
A:Accession: S12571
A:Molecule type: DNA
A:Residues: 1-59, 'I', 61-81, 83-348 <LOC2>
A:Cross-references: EMBL:X17606; NID:99785
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-348/Product: circumsporozoite protein #status predicted <MAT>
F:94-205/Region: 8-residue repeats
F:215-247/Region: 2-residue repeats
F:274-326/Domain: thrombospondin type 1 repeat homology <THRI>

17

Qy 8 RPSGGPSLD-----EKKIAKMEKASSVFNVNS 36
|||:::|:|:|:|:|:|:|:
Dp 300 RKGSNKAELTLIEDIDTEICKMDKCSSIFNVSN 334

RESULT 15
150739 gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
C:Species: Haplochromis burtoni
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C/Accession: 150739
R:White, S.A.; Kaaten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
A/Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
A/Reference number: 150739; MUID:95396797; PMID:7667296
A/Accession: 150739
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-98 <WHI>
A/Cross-references: EMBL:U31865; NID:905398; PIDN:AAC59691.1; PID:905399
C:Superfamily: gonadoliberin

2;

Qy 2 HWSYGLRPGSSGSLDE-----KKIAKMKKASSVF 31

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.58366 Seconds

(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGRFGSSGFSLEKXIAMEXASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	53.0	388	1	CSP_PLARE
2	91	49.7	397	1	CSP_PLAFO
3	91	49.7	412	1	CSP_PLAFA
4	91	49.7	424	1	CSP_PLAFT
5	90	49.2	442	1	CSP_PLAFW
6	61.5	33.6	61	1	GONI_SHEEP
7	59	32.2	63	1	GONI_MESAU
8	58	31.7	67	1	GONI_MACMU
9	58	31.7	89	1	GONI_XENLA
10	58	31.7	90	1	GONI_MOUSE
11	58	31.7	90	1	GONI_RANCA
12	58	31.7	91	1	GONI_PIG
13	58	31.7	92	1	GONI_HUMAN
14	58	31.7	92	1	GONI_RAT
15	58	31.7	92	1	GONI_TUPGB
16	58	31.7	339	1	CSP_PLABE
17	58	31.7	347	1	CSP_PLABA
18	56.5	30.9	90	1	GONI_DICLA
19	55.5	30.3	89	1	GONI_PORNO
20	55	30.1	94	1	GONI_HAPBU
21	55	30.1	367	1	CSP_PLAYO
22	55	30.1	721	1	THIC_SHEON
23	54	29.5	10	1	GONI_ALIMI
24	54	29.5	92	1	GONI_CHICK
25	53.5	29.2	90	1	GONI_HAPBU
26	52.5	28.7	485	1	NWSA_BACSU
27	52	28.4	97	1	GONI_MORSA
28	52	28.4	95	1	GONI_PAGMA
29	52	28.4	95	1	GONI_SPAU
30	52	28.4	99	1	GONI_DICLA
31	51.5	28.1	80	1	GONI_CLAGA
32	51.5	28.1	90	1	GONI_ORYIA
33	51.5	28.1	444	1	TIG_MYCGE

ALIGNMENTS

RESULT 1

CSP_PLARE STANDARD; PRT; 388 AA.

ID CSP_PLARE AC P26694;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Circumsporozoite protein precursor (CS).

OS Plasmodium reichenowi.

CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

CC NCBI_TaxID=5854;

CC [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=91201303; PubMed=2016283;

RL Lal A.A., Goldman I.F.;

RT "Circumsporozoite protein gene from Plasmodium reichenowi, a

chimpanzee malaria parasite evolutionarily related to the human

malaria parasite Plasmodium falciparum.";

RL J. Biol. Chem. 266:6686-6689(1991).

CC -!- FUNCTION: The circumsporozoite protein is the immunodominant

surface antigen on the sporozoite (the infective stage of the

malaria parasite that is transmitted from the mosquito to the

vertebrate host).

CC -!- MISCELLANEOUS: The C-terminal region is probably used for

anchoring the protein to the cell membrane. the repeat sequences

would be the surface antigen of the organism.

CC -!- SIMILARITY: Contains 1 TSP type-1 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M60972; AAA29561.1; -

CC PIR; A39756; A39756; Crcmsprzoite.

CC InterPro; IPR003067; Crcmsprzoite.

CC InterPro; IPR000884; TSP1.

CC Pfam; PF00090; tsp_1; 1.

CC PRINTS; PR01303; CRCMSPRZOITE.

CC SMART; SM00209; TSP1; 1.

CC PROSITE; PS00092; TSP1; 1.

CC Malaria; Sporozoite; Repeat; Signal.

CC SIGNAL; 1 16

FT CHAIN 17 388

FT DOMAIN 120 263

FT REPEAT 120 123

FT REPEAT 124 127

FT REPEAT 128 131

FT REPEAT 132 135

FT REPEAT 136 139

FT REPEAT 140 143

FT REPEAT 144 147

PROBABLE.
CIRCUMSPOROZOITE PROTEIN.
37 X 4 AA TANDEM REPEATS OF N-[AV]-[ND]-

```

FT REPEAT 148 151 8.
FT REPEAT 152 155 9.
FT REPEAT 156 159 10.
FT REPEAT 160 163 11.
FT REPEAT 164 167 12.
FT REPEAT 168 171 13.
FT REPEAT 172 175 14.
FT REPEAT 176 179 15.
FT REPEAT 180 183 16.
FT REPEAT 184 187 17.
FT REPEAT 188 191 18.
FT REPEAT 192 195 19.
FT REPEAT 196 199 20.
FT REPEAT 200 203 21.
FT REPEAT 204 207 22.
FT REPEAT 208 211 23.
FT REPEAT 212 215 24.
FT REPEAT 216 219 25.
FT REPEAT 220 223 26.
FT REPEAT 224 227 27.
FT REPEAT 228 231 28.
FT REPEAT 232 235 29.
FT REPEAT 236 239 30.
FT REPEAT 240 243 31.
FT REPEAT 244 247 32.
FT REPEAT 248 251 33.
FT REPEAT 252 255 34.
FT REPEAT 256 259 35.
FT REPEAT 260 263 36.
FT DOMAIN 313 366 TSP TYPE-1.
SQ SEQUENCE 388 AA; 42245 MW; C031EBFBEE2E35604 CRC64;

Query Match 53.0%; Score 97; DB 1; Length 388;
Best Local Similarity 57.9%; Pred. No. 43e-06;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPGGSPSLD-----EKKIAMEKASSVFNVNS 36
Db 337 IRPGSAGKPKDQDYNDLEKKIKMKCSSVFNVNS 374

RESULT 2
ID_CSP_PLAFO STANDARD; PRT; 397 AA.
AC P19597; Q25798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate vaccine antigen.";
RL Nucleic Acids Res. 17:5854-5854 (1989).
RN [2]
RP REVISIONS
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
RA Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human infectivity of a cloned line.";
RL Exp. Parasitol. 74:159-168 (1992).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=89364998; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate used in malaria vaccine trials.";
RL Mol. Biochem. Parasitol. 35:185-190 (1989).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant surface antigen on the sporozoite (the infective stage of the malaria parasite that is transmitted from the mosquito to the vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for anchoring the protein to the cell membrane. The repeat sequences would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; X15363; CAA33421.1; -
CC EMBL; M83886; AAA29521.1; -
CC EMBL; M22982; AAA29527.1; -
CC PIR; S05428; S05428;
CC InterPro; IPR003067; Crcmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZTOITE.
CC SMART; SMC0209; TSP1; 1.
CC PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 105 272 42 X 4 AA TANDEM REPEATS OF N-[AV] - [ND] - P.
FT REPEAT 105 108 1.
FT REPEAT 109 112 2.
FT REPEAT 113 116 3.
FT REPEAT 117 120 4.
FT REPEAT 121 124 5.
FT REPEAT 125 128 6.
FT REPEAT 129 132 7.
FT REPEAT 133 136 8.
FT REPEAT 137 140 9.
FT REPEAT 141 144 10.
FT REPEAT 145 148 11.
FT REPEAT 149 152 12.
FT REPEAT 153 156 13.
FT REPEAT 157 160 14.
FT REPEAT 161 164 15.
FT REPEAT 165 168 16.
FT REPEAT 169 172 17.
FT REPEAT 173 176 18.
FT REPEAT 177 180 19.
FT REPEAT 181 184 20.
FT REPEAT 185 188 21.
FT REPEAT 189 192 22.
FT REPEAT 193 196 23.
FT REPEAT 197 200 24.
FT REPEAT 201 204 25.
FT REPEAT 205 208 26.
FT REPEAT 209 212 27.
FT REPEAT 213 216 28.
FT REPEAT 217 220 29.
FT REPEAT 221 224 30.
FT REPEAT 225 228 31.
FT REPEAT 229 232 32.
FT REPEAT 233 236 33.
FT REPEAT 237 240 34.
FT REPEAT 241 244 35.

```

```

FT REPEAT 245 248 36.
FT REPEAT 249 252 37.
FT REPEAT 253 256 38.
FT REPEAT 257 260 39.
FT REPEAT 261 264 40.
FT REPEAT 265 268 41.
FT REPEAT 269 272 42.
FT DOMAIN 322 375 TSP TYPE-1.
FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9881146F59E3EA3 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 397;
Best Local Similarity 55.3%; Pred. No. 3.1e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKXIAMKAKSSVFNVNVS 36
DB 346 IKPGSANKPKDLDYNDIEKKICKMKCKSSVFNVNVS 393

RESULT 3
CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RA "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum."
RL Science 225:593-599(1984).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02194; AAA29524.1; -.
DR PIR; A03388; OZQOAF.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 412
FT DOMAIN 123 286
FT REPEAT 123 126 1.
FT REPEAT 127 130 2.
FT REPEAT 131 134 3.

```

```

FT REPEAT 135 138 4.
FT REPEAT 139 142 5.
FT REPEAT 143 146 6.
FT REPEAT 147 150 7.
FT REPEAT 151 154 8.
FT REPEAT 155 158 9.
FT REPEAT 159 162 10.
FT REPEAT 163 166 11.
FT REPEAT 167 170 12.
FT REPEAT 171 174 13.
FT REPEAT 175 178 14.
FT REPEAT 179 182 15.
FT REPEAT 183 186 16.
FT REPEAT 187 190 17.
FT REPEAT 191 194 18.
FT REPEAT 195 198 19.
FT REPEAT 199 202 20.
FT REPEAT 203 206 21.
FT REPEAT 207 210 22.
FT REPEAT 211 214 23.
FT REPEAT 215 218 24.
FT REPEAT 219 222 25.
FT REPEAT 223 226 26.
FT REPEAT 227 230 27.
FT REPEAT 231 234 28.
FT REPEAT 235 238 29.
FT REPEAT 239 242 30.
FT REPEAT 243 246 31.
FT REPEAT 247 250 32.
FT REPEAT 251 254 33.
FT REPEAT 255 258 34.
FT REPEAT 259 262 35.
FT REPEAT 263 266 36.
FT REPEAT 267 270 37.
FT REPEAT 271 274 38.
FT REPEAT 275 278 39.
FT REPEAT 279 282 40.
FT REPEAT 283 286 41.
FT DOMAIN 337 390 TSP TYPE-1.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE9065F8 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 412;
Best Local Similarity 55.3%; Pred. No. 3.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKXIAMKAKSSVFNVNVS 36
DB 361 IKPGSANKPKDLDYNDIEKKICKMKCKSSVFNVNVS 398

RESULT 4
CSP_PLAFT STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RA "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand."
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for

```



```
FT REPEAT 202 205 18.
FT REPEAT 206 209 19.
FT REPEAT 210 213 20.
FT REPEAT 211 217 21.
FT REPEAT 218 221 22.
FT REPEAT 222 225 23.
FT REPEAT 226 229 24.
FT REPEAT 230 233 25.
FT REPEAT 234 237 26.
FT REPEAT 238 241 27.
FT REPEAT 242 245 28.
FT REPEAT 246 249 29.
FT REPEAT 250 253 30.
FT REPEAT 254 257 31.
FT REPEAT 258 261 32.
FT REPEAT 262 265 33.
FT REPEAT 266 269 34.
FT REPEAT 270 273 35.
FT REPEAT 274 277 36.
FT REPEAT 278 281 37.
FT REPEAT 282 285 38.
FT REPEAT 286 289 39.
FT REPEAT 290 293 40.
FT REPEAT 294 297 41.
FT REPEAT 298 301 42.
FT REPEAT 302 305 43.
FT REPEAT 306 309 44.
FT REPEAT 310 313 45.
FT REPEAT 314 317 46.
FT DOMAIN 367 420 TSP TYPE-1.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B95E03 CRC64;

Query Match 49.2%; Score 90; DB 1; Length 442;
Best Local Similarity 55.3%; Pred. No. 4.9e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKIAKWEKASSVFNVNS 36
   :|||:|||||:|||||:|||||:|||||:|||||:
Db 391 IKFGSADPKOLDIYNDIEKKIKWEKCSSVFNVNS 428

RESULT 6
GONI_SHEEP
ID GONI_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment)
DE GNRH1 OR GNRH OR LHRH.
GN Ovis aries (Sheep).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN=Western ranger, TISSUE=Hypothalamus;
RA Rodriguez R.E.; Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgess R.; Burcher M.; Amos M.; Ling N.; Monahan M.; Rivier J.;
RA Fellows R.; Blackwell R.; Vale W.; Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
```

```
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02517; AAA03433.1; -.
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; GonadolibereinI.
CC Pfam; PF00446; GnRH; 1
CC PRINTS; PR01541; GONADOLIBERNI.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pyroglutamate carboxylic acid.
FT NON_TER 1 1
FT CHAIN 1 >61 PROGONADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYROGLUTAMATE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 33.6%; Score 61.5; DB 1; Length 61;
Best Local Similarity 42.9%; Pred. No. 0.061;
Matches 15; Conservative 3; Mismatches 2; Indels 15; Gaps 2;

OY 2 HWSYGLRPGSSGSPSLDEKIAKWEKASSVFNVNS 36
   |||||:|||||:|||||:|||||:|||||:
Db 2 HWSYGLRPGG-----KKNK-----NVIDS 21

RESULT 7
GONI_MESAU
ID GONI_MESAU STANDARD; PRT; 63 AA.
AC Q09183;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment)
DE GNRH1 OR GNRH OR LHRH.
GN Mesocricetus auratus (Golden hamster).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T.; Stevens P.J.; Zeitler P.; Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL; U91938; AAB51302.1; -.
CC DR InterPro; IPR002012; GnRH.
CC DR InterPro; IPR004079; GonadoliberinI.
CC DR Pfam; PF00446; GnRH; 1.
CC DR PRINTS; PR01541; GONADOLIBRN1.
CC DR PROSITE; PS00473; GnRH; 1.
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Pyrrolidone carboxylic acid.
CC FT NON_TER 1 1
CC FT CHAIN 1 >63
CC FT CHAIN 1 >63
CC FT PEPTIDE 1 10
CC FT PEPTIDE 14 >63
CC FT ACT_SITE 3 3
CC FT MOD_RES 1 1
CC FT MOD_RES 10 10
CC FT MOD_RES 63 63
CC FT NON_TER 63 63
CC FT SEQUENCE 63 AA; 7370 MW; FC9495676F77180 CRC64;
CC
CC Query Match 32.2%; Score 59; DB 1; Length 63;
CC Best Local Similarity 52.2%; Pred. No. 0.14;
CC Matches 12; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
CC
CC QY 2 HWSYGLRPGSS-----GPSLDE 18
CC DB 2 HWSYGLRPGKRNARLGDSPQE 24
CC
CC RESULT 8
CC GONI_MACMU STANDARD; PRT; 67 AA.
CC ID GONI_MACMU
CC AC P55247;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Progadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
CC DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
CC DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
CC DE (Fragment).
CC GN GNRH1 OR GNRH OR LHRH.
CC OS Macaca mulatta (Rhesus macaque).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC OC Cercopithecoidea; Macaca.
CC OX NCBI_TaxID=9544;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Hypothalamus;
CC RX MEDLINE=95124501; PubMed=7545971;
CC RA Ma Y.-J., Costa M.E., Ojeda S.R.;
CC RT "Developmental expression of the genes encoding transforming growth
CC RT factor alpha and its receptor in the hypothalamus of female rhesus
CC RT macaques."
CC RL Neuroendocrinology 60:346-359(1994).
CC CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC CC the secretion of both luteinizing and follicle-stimulating
CC CC hormones.
CC CC -!- SUBCELLULAR LOCATION: Secreted.
CC CC -!- SIMILARITY: Belongs to the GnRH family.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL; S75918; AAB33096.1; -.
CC DR PIR; I78541; I78541.
CC DR InterPro; IPR002012; GnRH.
CC DR InterPro; IPR004079; GonadoliberinI.
CC DR Pfam; PF00446; GnRH; 1.
CC DR PRINTS; PR01541; GONADOLIBRN1.
CC DR PROSITE; PS00473; GnRH; 1.
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Pyrrolidone carboxylic acid.
CC FT NON_TER 1 1
CC FT SIGNAL <1 5
CC FT CHAIN 1 >67
CC FT CHAIN 1 >67
CC FT PEPTIDE 19 >67
CC FT PEPTIDE 19 >67
CC FT ACT_SITE 8 8
CC FT MOD_RES 6 6
CC FT MOD_RES 15 15
CC FT MOD_RES 15 15
CC FT NON_TER 67 67
CC FT NON_TER 67 67
CC FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
CC
CC Query Match 31.7%; Score 58; DB 1; Length 67;
CC Best Local Similarity 100.0%; Pred. No. 0.21;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 HWSYGLRPG 10
CC DB 7 HWSYGLRPG 15
CC
CC RESULT 9
CC GONI_XENLA STANDARD; PRT; 89 AA.
CC ID GONI_XENLA
CC AC P45656;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
CC DE (LH-RH) (Luliberin I).
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
CC OC Xenopodinae; Xenopus.
CC OX NCBI_TaxID=8355;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Forebrain;
CC RX MEDLINE=94185563; PubMed=8137750;
CC RA Hayes W.P., Wray S., Battey J.F.;
CC RT "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
CC RT mammalian-like expression pattern and conserved domains in
CC RT GnRH-associated peptide, but brain onset is delayed until
CC RT metamorphosis."
CC RL Endocrinology 134:1835-1844(1994).
CC CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC CC -!- SUBCELLULAR LOCATION: Secreted.
CC CC -!- SIMILARITY: Belongs to the GnRH family.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERIN I.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 89
 FT PEPTIDE 24 89
 FT PEPTIDE 37 89
 FT PEPTIDE 37 89
 FT PEPTIDE 37 89
 FT MOD_RES 24 24
 FT MOD_RES 33 33
 FT MOD_RES 33 33
 SQ SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;
 Query Match 31.7%; Score 58; DB 1; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 25 HWSYGLRPG 33
 RESULT 10
 GON1_MOUSE STANDARD; PRT; 90 AA.
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 hormone I) (Gnrh I) (Luliberin I); Prolactin release-inhibiting factor
 I]
 DE GNRH1 OR GNRH.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97069928; PubMed=3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolic K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371 (1986).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 the secretion of both luteinizing and follicle-stimulating
 hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the Gnrh family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; W14872; AAA37717.1;
 DR PIR; A47578; RMSG.
 DR MGD; MG1:95789; Gnrh.
 DR InterPro; IPR002012; Gnrh.
 DR InterPro; IPR004079; GonadolibirinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERIN I.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 FT Placenta; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 21

FT CHAIN 22 90
 FT PEPTIDE 22 31
 FT PEPTIDE 35 90
 FT ACT_SITE 24 24
 FT ACT_SITE 24 24
 FT ACT_SITE 24 24
 FT MOD_RES 22 22
 FT MOD_RES 31 31
 FT MOD_RES 31 31
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA826E4D9 CRC64;
 Query Match 31.7%; Score 58; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 23 HWSYGLRPG 31
 RESULT 11
 GON1_RANCA STANDARD; PRT; 90 AA.
 AC Q90Y63;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LHRH I)
 (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I (GAP1)].
 DE GNRH1 OR GNRH.
 GN Rana catesbeiana (Bull frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Forebrain; PubMed=1170016;
 RX MEDLINE=21102951;
 RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
 RA Kwon H.B.;
 RT "Cloning and characterization of cDNAs encoding the Gnrh1 and Gnrh2
 precursors from bullfrog (Rana catesbeiana).";
 RL J. Exp. Zool. 289:190-201 (2001).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
 similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Forebrain.
 CC -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
 during post-breeding. Not expressed in pituitary.
 CC -!- SIMILARITY: Belongs to the Gnrh family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF188754; AA05972.1;
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
 DR GO; GO:0009755; P:hormone mediated signaling; NAS.
 DR GO; GO:0000003; P:reproduction; NAS.
 DR InterPro; IPR002012; Gnrh.
 DR InterPro; IPR004079; GonadolibirinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERIN I.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Signal;
 Pyroglutamate carboxylic acid.
 FT SIGNAL 1 24
 FT CHAIN 25 90
 FT PEPTIDE 25 34
 FT PEPTIDE 25 34


```
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [6]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373 (1999).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutrepulse or Lutrelaf (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01059; CAA25526.1; -
DR EMBL; M12578; AAA35916.1; -
DR EMBL; X15215; CAA33285.1; -
DR PIR; S05308; RHUUG.
DR Genew; HGNC:4419; GNRH1.
DR MM; 152760; -
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007185; P:signal transduction; TAS.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH.
DR PRINTS; PR01541; GNADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal; Polymorphism;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT VARIANT 16 16
FT SIGNAL 1 23
FT SEQUENCE 92 AA; 10380 MW; 30A72221E076FA79 CRC64;
Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred.No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
```

```
RESULT 14
GONI_RAT
AC GONI_RAT STANDARD; PRT; 92 AA.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RT Mol. Endocrinol. 3:1257-1262 (1989).
RL [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262 (1989).
RP [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454 (1992).
RP [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517 (1987).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Central nervous system.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RHRTG.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GNADOLIBRN1.
```

DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 15
GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN GNRH1 OR GNRH.
OS Tupala glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=37347;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U63326; AAB16837.1; --
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; Gonadoliberin I.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 24 24 ACTIVITY.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64; SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64; AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
Search completed: March 10, 2004, 09:13:57
Job time : 6.58366 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 08:58:54 ; Search time 35.5798 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-16
Perfect score: 183
Sequence: 1 XHWSYGLRPGSSGSLDEKKIAXMEKASSVFNVNS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriaph.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	55.7	117	5	Q25797 Plasmodium
2	102	55.7	408	5	Q25729 Plasmodium
3	102	55.7	436	5	Q27325 Plasmodium
4	98	53.6	79	5	Q900Q2 Plasmodium
5	98	53.6	392	5	Q819H8 Plasmodium
6	98	53.6	396	5	Q819I1 Plasmodium
7	98	53.6	396	5	Q819I0 Plasmodium
8	98	53.6	396	5	Q819H9 Plasmodium
9	97	53.0	360	5	Q819H7 Plasmodium
10	97	53.0	360	5	Q819H6 Plasmodium
11	97	53.0	412	5	Q819H5 Plasmodium
12	97	53.0	420	5	Q25838 Plasmodium
13	94	51.4	389	5	Q819J2 Plasmodium
14	94	51.4	393	5	Q819J1 Plasmodium
15	91	49.7	69	5	Q8MZK7 Plasmodium
16	91	49.7	80	5	Q9U0P3 Plasmodium

17	91	49.7	80	5	Q9U0P7 Plasmodium
18	91	49.7	80	5	Q8TW97 Plasmodium
19	91	49.7	80	5	Q9U0P1 Plasmodium
20	91	49.7	80	5	Q9TVQ0 Plasmodium
21	91	49.7	80	5	Q9U0P9 Plasmodium
22	91	49.7	80	5	Q8TVP9 Plasmodium
23	91	49.7	80	5	Q9U0Q0 Plasmodium
24	91	49.7	80	5	Q9U0P5 Plasmodium
25	91	49.7	80	5	Q8TW76 Plasmodium
26	91	49.7	80	5	Q8U0Q4 Plasmodium
27	91	49.7	80	5	Q9U0P6 Plasmodium
28	91	49.7	80	5	Q9U0P2 Plasmodium
29	91	49.7	80	5	Q9U0P8 Plasmodium
30	91	49.7	80	5	Q9U0P4 Plasmodium
31	91	49.7	80	5	Q9TVN9 Plasmodium
32	91	49.7	80	5	Q8TW83 Plasmodium
33	91	49.7	115	5	Q25839 Plasmodium
34	91	49.7	115	5	Q25835 Plasmodium
35	91	49.7	115	5	Q25836 Plasmodium
36	91	49.7	115	5	Q9TVW7 Plasmodium
37	91	49.7	117	5	Q25795 Plasmodium
38	91	49.7	117	5	Q25796 Plasmodium
39	91	49.7	117	5	Q25794 Plasmodium
40	91	49.7	356	5	Q819G5 Plasmodium
41	91	49.7	363	5	Q819H3 Plasmodium
42	91	49.7	364	5	Q819G4 Plasmodium
43	91	49.7	368	5	Q819G3 Plasmodium
44	91	49.7	380	5	Q819G8 Plasmodium
45	91	49.7	380	5	Q819G1 Plasmodium

ALIGNMENTS

RESULT 1

Q25797 PRELIMINARY; PRT; 117 AA.
AC Q25797;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doolan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
RT circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M77205; AAA29519.2;
DR GO: GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13043 MW; 0DA711D86C0B03C1 CRC64;

Query Match 55.7%; Score 102; DB 5; Length 117;
Best Local Similarity 60.5%; Pred. No. 1.1e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

Qy 7 LRGSSGSLD-----EKKIAXMEKASSVFNVNS 36
:|||||
Db 66 IKPGSAGKNDLNDYENDIEKICKMEKCSVFNVNS 103
:|||||

RESULT 2

Q25729 ID Q25729 PRELIMINARY; PRT; 408 AA.
AC Q25729;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Santa Lucia;
RA Qari S.H., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20969; AAA63153.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria.
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;
Query Match 55.7%; Score 102; DB 5; Length 408;
Best Local Similarity 60.5%; Pred. No. 4.3e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;
QY 7 LRPGSSGSPSLD-----EKIKAKMEKASSVFNVNS 36
DB 357 IKFGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNS 394
RESULT 3
Q27325 ID Q27325 PRELIMINARY; PRT; 436 AA.
AC Q27325;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wittz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Jongwutiwes S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83164; AAA29542.1; -
DR EMBL; M83150; AAA29563.1; -
DR EMBL; M83163; AAA29576.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.

DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 436 AA; 46688 MW; 5B42FF3348B68655 CRC64;
Query Match 55.7%; Score 102; DB 5; Length 436;
Best Local Similarity 60.5%; Pred. No. 4.6e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;
QY 7 LRPGSSGSPSLD-----EKIKAKMEKASSVFNVNS 36
DB 385 IKFGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNS 422
RESULT 4
Q9U0Q2 ID Q9U0Q2 PRELIMINARY; PRT; 79 AA.
AC Q9U0Q2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS1;
RA de Stricker K., Vuust J., Jeppen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269945; CAB64243.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1 79
SQ SEQUENCE 79 AA; 8893 MW; 56FBFA76D859B416 CRC64;
Query Match 53.6%; Score 98; DB 5; Length 79;
Best Local Similarity 59.5%; Pred. No. 2.6e-06;
Matches 22; Conservative 3; Mismatches 4; Indels 8; Gaps 1;
QY 7 LRPGSSGSPSLD-----EKIKAKMEKASSVFNVNS 35
DB 43 IKFGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNS 79
RESULT 5
Q819H8 ID Q819H8 PRELIMINARY; PRT; 392 AA.
AC Q819H8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asambo Bay;
RX MEDLINE=22356746; PubMed=12467976;

```

RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540461; AAN87595.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 392 AA; 42385 MW; 99DAAD1629801E0C CRC64;

Query Match 53.6%; Score 98; DB 5; Length 392;
Best Local Similarity 63.2%; Pred. No. 1.6e-05;
Matches 24; Conservative 3; Mismatches 3; Indels 8; Gaps 2;

Qy 7 LRPGSSGPS---LD-----EKKIAMEKASSVFNVVNS 36
Db 341 IKPGSAGSKNELDYENDIEKKICKMEKCSSVFNVVNS 378

RESULT 6
Q81911 ID Q81911 PRELIMINARY; PRT; 396 AA.
AC Q81911;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540458; AAN87592.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42781 MW; 97070A9ED8D517D1 CRC64;

Query Match 53.6%; Score 98; DB 5; Length 396;
Best Local Similarity 63.2%; Pred. No. 1.6e-05;
Matches 24; Conservative 3; Mismatches 3; Indels 8; Gaps 2;

Qy 7 LRPGSSGPS---LD-----EKKIAMEKASSVFNVVNS 36
Db 345 IKPGSAGSKNELDYENDIEKKICKMEKCSSVFNVVNS 382

RESULT 7
Q81910 ID Q81910 PRELIMINARY; PRT; 396 AA.
AC Q81910;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

```

```

DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540459; AAN87593.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42782 MW; 87564F9AD8D517D1 CRC64;

Query Match 53.6%; Score 98; DB 5; Length 396;
Best Local Similarity 63.2%; Pred. No. 1.6e-05;
Matches 24; Conservative 3; Mismatches 3; Indels 8; Gaps 2;

Qy 7 LRPGSSGPS---LD-----EKKIAMEKASSVFNVVNS 36
Db 345 IKPGSAGSKNELDYENDIEKKICKMEKCSSVFNVVNS 382

RESULT 8
Q819H9 ID Q819H9 PRELIMINARY; PRT; 396 AA.
AC Q819H9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540460; AAN87594.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42768 MW; 709FA806690FD17C CRC64;

Query Match 53.6%; Score 98; DB 5; Length 396;
Best Local Similarity 63.2%; Pred. No. 1.6e-05;
Matches 24; Conservative 3; Mismatches 3; Indels 8; Gaps 2;

Qy 7 LRPGSSGPS---LD-----EKKIAMEKASSVFNVVNS 36
Db 345 IKPGSAGSKNELDYENDIEKKICKMEKCSSVFNVVNS 382

```

```
SQ SEQUENCE 360 AA; 39062 MW; 65058844270D666C CRC64;
Query Match 53.0%; Score 97; DB 5; Length 360;
Best Local Similarity 57.9%; Pred. No. 2e-05; Mismatches 3; Indels 8; Gaps 1;
Matches 22; Conservative

QY 7 LRPSSGSPSLD-----EKKIAYKEKASSVFNVNS 36
      ::|||:|
Db 309 IRPGSAGKPKDELVDYANDIEKKICKMEKCSSVFNVNS 346
      ::|||:|

RESULT 11
Q819H5 PRELIMINARY; PRT; 412 AA.
ID Q819H5;
AC Q819H5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Asebo Bay; PubMed=12467976;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asebo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540462; AAN87596.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
SQ SEQUENCE 360 AA; 39093 MW; 6504CC012649236C CRC64;

Query Match 53.0%; Score 97; DB 5; Length 360;
Best Local Similarity 57.9%; Pred. No. 2e-05;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAYKEKASSVFNVNS 36
      ::|||:|
Db 309 IRPGSAGKPKDELVDYANDIEKKICKMEKCSSVFNVNS 346
      ::|||:|

RESULT 10
Q819H6 PRELIMINARY; PRT; 360 AA.
ID Q819H6;
AC Q819H6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Asebo Bay; PubMed=12467976;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asebo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540463; AAN87597.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
SQ SEQUENCE 360 AA; 39093 MW; 6504CC012649236C CRC64;

Query Match 53.0%; Score 97; DB 5; Length 412;
Best Local Similarity 57.9%; Pred. No. 2.3e-05;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAYKEKASSVFNVNS 36
      ::|||:|
Db 361 IRPGSAGKPKDELVDYANDIEKKICKMEKCSSVFNVNS 398
      ::|||:|

RESULT 12
Q25838 PRELIMINARY; PRT; 420 AA.
ID Q25838;
AC Q25838;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=835b; PubMed=7985759;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwattives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83161; AAA29574.1; -
```

```
DR GO: GO:0003793; F: defense/immunity protein activity; IEA.
DR InterPro: IPR003067; C: Crcmsprzoite.
DR Pfam: PF00090; tsep 1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR SEQUENCE 420 AA; 45155 MW; 3A85B92432C3893C CRC64;

Query Match 53.0%; Score 97; DB 5; Length 420;
Best Local Similarity 57.9%; Pred. No. 2.3e-05;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 369 IKPGSAGKPKDELVDYENDIEKKICKMEKCSSVFNVNS 406

RESULT 13
Q819J2 PRELIMINARY; PRT; 389 AA.
AC Q819J2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1];
RP STRAIN=Asambo Bay;
RC MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asambo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL: AF540443; AAN87577.1; -
DR GO: GO:0003793; F: defense/immunity protein activity; IEA.
DR InterPro: IPR003067; C: Crcmsprzoite.
DR Pfam: PF00090; tsep 1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR SEQUENCE 389 AA; 41880 MW; 4F2C32A159B13F20 CRC64;

Query Match 51.4%; Score 94; DB 5; Length 389;
Best Local Similarity 60.5%; Pred. No. 5.8e-05;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 338 IKPGSAGKPKDELVDYENDIEKKICKMEKCSSVFNVNS 375

RESULT 14
Q819J1 PRELIMINARY; PRT; 393 AA.
AC Q819J1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Asambo Bay;
RC MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asambo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL: AF540443; AAN87577.1; -
DR GO: GO:0003793; F: defense/immunity protein activity; IEA.
DR InterPro: IPR003067; C: Crcmsprzoite.
DR Pfam: PF00090; tsep 1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR SEQUENCE 393 AA; 42276 MW; 7D2C7CFAA3153A9 CRC64;

Query Match 51.4%; Score 94; DB 5; Length 393;
Best Local Similarity 60.5%; Pred. No. 5.8e-05;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 342 IKPGSAGKPKDELVDYENDIEKKICKMEKCSSVFNVNS 379

RESULT 15
Q8MZK7 PRELIMINARY; PRT; 69 AA.
AC Q8MZK7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein II (Fragment).
GN CSPII.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=FCCI/HN;
RA Chen H.H., Yu X.B., Wu Z.D., Xu J.;
RT "Gene cloning of circumsporozoite protein (CSP) II gene from
RT Plasmodium falciparum (FCCI/HN).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY093572; AAM19072.1; -
DR GO: GO:0003793; F: defense/immunity protein activity; IEA.
DR InterPro: IPR003067; C: Crcmsprzoite.
DR Pfam: PF00090; tsep 1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR NON_TER 1;
DR SEQUENCE 69 AA; 7670 MW; F3E1AA54671FE6FF CRC64;

Query Match 49.7%; Score 91; DB 5; Length 69;
Best Local Similarity 55.3%; Pred. No. 2.2e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 18 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 55

Search completed: March 10, 2004, 09:25:38
Job time : 35:5798 secs
```

```
STRAIN=Yaounde;
MEDLINE=22356746; PubMed=12467976;
Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
Magris M., Biswas S., Kariuki S., Lal A.A.;
"A study of genetic diversity in the gene encoding the
circumsporozoite protein (CSP) of Plasmodium falciparum from different
transmission areas-XVI. Asambo Bay Cohort Project.";
Mol. Biochem. Parasitol. 125:83-90(2002).
EMBL: AF540444; AAN87578.1; -
GO: GO:0003793; F: defense/immunity protein activity; IEA.
InterPro: IPR003067; C: Crcmsprzoite.
InterPro: IPR000884; TSP1.
Pfam: PF00090; tsep 1; 1.
PRINTS: PR01303; CRCMSPRZOITE.
SMART: SM00209; TSP1; 1.
PROSITE: PS50092; TSP1; 1.
SEQUENCE 393 AA; 42276 MW; 7D2C7CFAA3153A9 CRC64;

Query Match 51.4%; Score 94; DB 5; Length 393;
Best Local Similarity 60.5%; Pred. No. 5.8e-05;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 342 IKPGSAGKPKDELVDYENDIEKKICKMEKCSSVFNVNS 379

RESULT 15
Q8MZK7 PRELIMINARY; PRT; 69 AA.
AC Q8MZK7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein II (Fragment).
GN CSPII.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=FCCI/HN;
RA Chen H.H., Yu X.B., Wu Z.D., Xu J.;
RT "Gene cloning of circumsporozoite protein (CSP) II gene from
RT Plasmodium falciparum (FCCI/HN).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY093572; AAM19072.1; -
DR GO: GO:0003793; F: defense/immunity protein activity; IEA.
DR InterPro: IPR003067; C: Crcmsprzoite.
DR Pfam: PF00090; tsep 1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR NON_TER 1;
DR SEQUENCE 69 AA; 7670 MW; F3E1AA54671FE6FF CRC64;

Query Match 49.7%; Score 91; DB 5; Length 69;
Best Local Similarity 55.3%; Pred. No. 2.2e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 18 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 55

Search completed: March 10, 2004, 09:25:38
Job time : 35:5798 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 08:58:48 ; Search time 54,2101 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-16
Perfect score: 183
Sequence: 1 XHWSYGLRPGSSGSLDEKKIAKXKASSVFNVNS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	99.5	36	5 AAU11427	AAU11427 Synthetic
2	182	99.5	51	5 AAU11431	AAU11431 Synthetic
3	95	51.9	20	5 AAU11414	AAU11414 P. falcip
4	95	51.9	33	5 AAU11423	AAU11423 Synthetic
5	91	49.7	309	2 AAR13175	AAR13175 NS1_81-RL
6	91	49.7	319	2 AAR07945	AAR07945 NS181RLFA
7	91	49.7	319	2 AAR13176	AAR13176 NS1_81-RL
8	91	49.7	327	2 AAR13177	AAR13177 NS1_81-RL
9	91	49.7	335	2 AAR13178	AAR13178 NS1_81-RL
10	91	49.7	335	2 AAR13179	AAR13179 NS1_81(NV
11	91	49.7	396	7 ABO23530	ABO23530 Plasmodiu
12	91	49.7	411	1 AAP83144	AP83144 Sequence
13	91	49.7	412	1 AAP60416	AP60416 CS protei
14	91	49.7	424	2 AAR37797	AAR37797 RTS* prot
15	91	49.7	424	2 AAR37796	AAR37796 RTS prote
16	89.5	48.9	33	4 AAG63663	AAG63663 Peptide c
17	89.5	48.9	33	4 AAG63516	AAG63516 A peptide
18	89	48.6	19	4 AAM98951	AAM98951 Vaccine r
19	89	48.6	21	1 AAP91504	AP91504 Sequence
20	89	48.6	21	2 AAR78920	AAR78920 Malaria c
21	89	48.6	21	2 AAR75955	AAR75955 P. falcip
22	89	48.6	21	2 AAR70912	AAR70912 Malaria c
23	89	48.6	21	2 AAR82586	AAR82586 Plasmodiu
24	89	48.6	21	2 AAW05612	AAW05612 Circumpo
25	89	48.6	21	2 AAW35440	AAW35440 T-cell st

26	89	48.6	21	2 AAY23252	AAY23252 Peptide d
27	89	48.6	21	3 AAY58777	AAY58777 Unidentif
28	89	48.6	21	3 AAY80071	AAY80071 Pathogen
29	89	48.6	21	3 AAY54553	AAY54553 T helper
30	89	48.6	21	4 AAB98457	AAB98457 Plasmodiu
31	89	48.6	21	4 AAB84447	AAB84447 Sequence
32	89	48.6	21	4 AAG84517	AAG84517 Plasmodiu
33	89	48.6	21	4 AAB99706	AAB99706 Plasmodiu
34	89	48.6	21	4 AAG88269	AAG88269 Plasmodiu
35	89	48.6	21	4 AAG62428	AAG62428 Plasmodiu
36	89	48.6	21	4 AAG89366	AAG89366 Plasmodiu
37	89	48.6	21	5 ABB94469	ABB94469 Plasmodiu
38	89	48.6	21	5 ABB78050	ABB78050 Loosely M
39	89	48.6	21	5 ABJ11372	ABJ11372 P falcipa
40	89	48.6	21	5 ABJ05781	ABJ05781 P falcipa
41	89	48.6	21	5 ABP51501	ABP51501 Malaria c
42	89	48.6	21	5 AAU95369	AAU95369 Plasmodiu
43	89	48.6	21	5 AEG34858	AEG34858 P. falcip
44	89	48.6	21	5 ABJ01952	ABJ01952 158P1D7 r
45	89	48.6	21	5 AAU91545	AAU91545 P. falcip

ALIGNMENTS

RESULT 1
AAU11427
ID AAU11427 standard; peptide; 36 AA.
XX AAU11427;
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 8.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1..10
FT Misc-difference 1 /note= "Gonadotrophin releasing hormone epitope"
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11..16
FT /note= "Spacer peptide"
FT Peptide 17..36
FT /note= "Malaria CSP protein (378-398 aa)"
PN WO200185763-A2.
BD 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
PF
XX 05-MAY-2000; 2000US-0202328P.
PR
XX (APHT-) APHTON CORP.
PA
XX Grimes S, Michaeli D, Stevens VC;
PI WPI; 2002-049440/06.
DR
XX Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having

DR WPI

XX

XX Novel synthetic immunogen for inducing immune response against
 FT Gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX Disclosure; Page 28; 43pp; English.
 PS
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a peptide
 CC from circumsporozoite protein, CSP, a promiscuous helper T-cell peptide
 CC epitope used in the immunogen of the invention
 XX Sequence 20 AA;
 SQ
 Query Match 51.9%; Score 95; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 DEKTIARKEKASSVFNVNS 36
 DB 1 DEKTIARKEKASSVFNVNS 20
 RESULT 4
 ID AAU11423 standard; peptide; 33 AA.
 AC AAU11423;
 DT 12-MAR-2002 (first entry)
 XX Synthetic immunogen peptide 4.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Plasmodium falciparum.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 FT Peptide 1..20
 FT Peptide /note= "Malaria CSP protein (378-398 aa)"
 FT Peptide 21..24
 FT Peptide /note= "Spacer peptide"
 FT Peptide 25..33
 FT Modified-site 33
 FT /note= "Amidated glycine or glycine amide"
 XX WO200185763-A2.
 PN 15-NOV-2001.
 PD 04-MAY-2001; 2001WO-US014363.
 PF 05-MAY-2000; 2000US-0202328P.
 PR (APPT-) APHTON CORP.
 PA

XX Grimes S, Michaeli D, Stevens VC;
 PI WPI; 2002-049440/06.
 DR Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX Claim 11; Page 8; 43pp; English.
 PS
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX Sequence 33 AA;
 SQ
 Query Match 51.9%; Score 95; DB 5; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 DEKTIARKEKASSVFNVNS 36
 DB 1 DEKTIARKEKASSVFNVNS 20
 RESULT 5
 ID AAR13175 standard; protein; 309 AA.
 AC AAR13175;
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 29-AUG-1991 (first entry)
 XX NS1_81-RLfdelta9.
 DE Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
 KW influenza virus; non-structural protein 1; fusion.
 XX Plasmodium falciparum.
 OS Influenza virus; (A/PR/8/34/).
 XX Key Location/Qualifiers
 FT Region 1..81
 FT /label= N-terminal of NS1
 FT /note= "Influenza virus nonstructural protein 1"
 FT Peptide 82..87
 FT /label= synthetic linker
 FT Region 88
 FT /label= artifact
 FT /note= "see comments"
 FT Region 89..193
 FT /label= AAs 19-123 of CS protein
 FT /note= "Region 1 contg. flanking region less signal
 FT sequence"
 FT 194..309
 FT /label= AAs 297-412 of CS protein
 FT /note= "Region II flanking region minus 9 N-term- inal
 FT AAs"
 XX BP432965-A.

```

XX PD 19-JUN-1991.
XX PF 06-DEC-1990; 90EP-00313257.
XX PR 08-DEC-1989; 89US-00447746.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PI (USSA ) US SEC OF ARMY.
XX PR (BIOM-) BIOMEDICAL RES INST.
XX PA (GROS/) GROSS M S.
XX PI Gross MS, Gordon DM, Hollingdal MR;
XX WPI; 1991-179771/25.
XX PR Polypeptide comprising immunogenic determinants from P falciparum - for
XX PT vaccine against malaria infection in humans.
XX PS Example 1; Page 7; 18pp; English.
XX CC The polypeptide is prepd. by genetic engineering of genes encoding the P.
XX CC falciparum circumsporozoite (CS) protein [Dame et al., Science 225 : 593
XX CC (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et
XX CC al., Nucleic Acids Research, 8 : 5845 (1980)]. The DNA encoding the 1st
XX CC 81 AAs of the N-terminal of NS1 (NS1 81) is linked via a synthetic
XX CC sequence to DNA encoding Region I contg. flanking regionless the 18 AA
XX CC signal region, which in turn is fused to DNA encoding Region II-contg.
XX CC flanking region less the first nine N-terminal AAs. This CS fusion is
XX CC designated Rlfelta9. The Pro residue separating the Asp (at the C-
XX CC terminal of the linker) from Rlfelta9 is an artifact of a filled in
XX CC BamHI site. The peptide can be used in a vaccine for protection against
XX CC malaria. See also AAR12306-R12311 and AAR13176-R13179. (Updated on 25-MAR
XX CC -2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX CC field)
XX SQ Sequence 309 AA;
    Query Match 49.7%; Score 91; DB 2; Length 309;
    Best Local Similarity 55.3%; Pred. NO. 0.00011;
    Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
QY 7 LRPGSGPSLD-----EKKIAQMEKASSVFNVNS 36
    :|||:|
DB 258 IKPGSANKPKDLDYENDIEKKICRMKCKSVFNVNS 295

RESULT 6
AAR07945
ID AAR07945 standard; protein; 319 AA.
AC AAR07945;
XX 25-MAR-2003 (revised)
DT 22-FEB-1991 (first entry)
XX NS181RLFAuth plasmid product.
XX Malaria; vaccine.
XX Plasmodium falciparum.
XX Key Location/Qualifiers
XX Domain 1..81
XX /label= NS181 protein fragment
XX /note= "from plasmid pMG-1"
XX Domain 89..193
XX /label= Fragment of circumsporozoite protein
XX /note= "204..319"
XX Domain 204..319
XX /label= Fragment of circumsporozoite protein
XX EP398540-A.
XX

PD 22-NOV-1990.
XX 01-MAY-1990; 90EP-00304720.
XX 03-MAY-1989; 89US-00346863.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (GROS/) GROSS M S.
XX Gross MS, Young JF;
XX WPI; 1990-350299/47.
XX N-PSDB; AAQ06580.
XX New polypeptide used in malaria vaccine - comprises immunogenic
XX determinants from 1st and 2nd flanking regions of plasmodium surface
XX protein and intermediate repeat domain.
XX PS Example 2; Page 11-12; 24pp; English.
XX CC The product is useful in preparation of vaccines for treatment and
XX CC prophylaxis of plasmodium sporozite infection. It may be easily produced
XX CC in large pure quantities from a transformed E.coli expression system.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 319 AA;
    Query Match 49.7%; Score 91; DB 2; Length 319;
    Best Local Similarity 55.3%; Pred. NO. 0.00012;
    Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
QY 7 LRPGSGPSLD-----EKKIAQMEKASSVFNVNS 36
    :|||:|
DB 268 IKPGSANKPKDLDYENDIEKKICRMKCKSVFNVNS 305

RESULT 7
AAR13176
ID AAR13176 standard; protein; 319 AA.
XX AAR13176;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 29-AUG-1991 (first entry)
XX NS1_81-RLfAuth.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
XX influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
XX Influenza virus; (A/PR/8/34).
XX Key Location/Qualifiers
XX Region 1..81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Peptide 82..87
XX /label= synthetic linker
XX Region 88
XX /label= artifact
XX /note= "see comments"
XX Region 89..193
XX /label= AAs 19-123 of CS protein
XX /note= "Region 1 contg. flanking region less signal
XX sequence"
XX Region 194
XX /label= artifact
XX /note= "see comments"
XX Region 195..319
XX /label= AAs 288-412 of CS protein
XX /note= "Region II flanking region"
XX FT

```



```
AC AAR13178;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 29-AUG-1991 (first entry)
XX
XX NS1_81 (NANP)4RLfAuth.
XX
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
KW influenza virus; non-structural protein 1; fusion.
XX
XX Plasmodium falciparum.
XX
XX Influenza virus; (A/PR/8/34).
XX
XX Key Location/Qualifiers
XX 1. .81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX
XX 82. .97
XX /label= immunodominant repeat region
XX /note= "four tetrapeptide repeat units"
XX
XX 98. .103
XX /label= synthetic linker
XX
XX 104
XX /label= artifact
XX /note= "see comments"
XX
XX 105. .209
XX /label= AAs 19-123 of CS protein
XX /note= "Region 1 contg. flanking region less signal
XX
XX 210
XX /label= artifact
XX /note= "see comments"
XX
XX 211. .335
XX /label= AAs 288-412 of CS protein
XX /note= "Region II flanking region"
XX
XX EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-00313257.
XX
XX 08-DEC-1989; 89US-00447746.
XX
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROS/) GROSS M S.
XX
XX Gross MS, Gordon DM, Hollingdal MR;
XX WPI; 1991-179771/25.
XX
XX Polypeptide comprising immunogenic determinants from P falciparum - for
XX vaccine against malaria infection in humans.
XX
XX Example 4; Page 11; 18pp; English.
XX
XX The polypeptide is prep'd. by genetic engineering of genes encoding the P.
XX falciparum circumsporozoite (CS) protein [Dane et al., Science 225 : 593
XX (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et
XX al., Nucleic Acids Research, 8 : 5845 (1980)]. The DNA encoding the 1st
XX 81 AAs of the N-terminal of NS1 (NS1_81) is linked to a synthetic
XX sequence encoding four repeat units from the immunodominant region, which
XX in turn is linked via a synthetic sequence to DNA encoding Region I
XX contg. flanking region less the 18 AA signal region. This is linked to
XX DNA encoding Region II-contg. flanking region. The Pro residue sep-
XX arating the Asp (at the C-terminal of the linker) from the Region I-
XX contg. CS flanking region is an artifact of a filled-in BamHI site; the
XX Gly separating the Region I and II-contg. CS flanking regions is an
XX artifact of a synthetic FokI/TthII I linker. The peptide can be used in
XX a vaccine for protection against malaria. See also AAR12306-R12311 and
```

```
CC AAR13175-R13179. (Updated on 25-MAR-2003 to correct PA field.) (Updated
CC on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 335 AA;
XX
XX Query Match 49.7%; Score 91; DB 2; Length 335;
XX Best Local Similarity 55.3%; Pred. NO. 0.00013;
XX Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
XX
XX 7 LREGSGPSLD-----EKXIAXMEKASSVFNVNS 36
XX :||||:|||||
XX 284 IREGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 321
XX
XX RESULT 10
XX AAR13179
XX ID AAR13179 standard; protein; 335 AA.
XX
XX AAR13179;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 29-AUG-1991 (first entry)
XX
XX NS1_81 (NVPD)4RLfAuth.
XX
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
KW influenza virus; non-structural protein 1; fusion.
XX
XX Plasmodium falciparum.
XX
XX Influenza virus; (A/PR/8/34).
XX
XX Key Location/Qualifiers
XX 1. .81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX
XX 82. .97
XX /label= immunodominant repeat region
XX /note= "four variant tetrapeptide repeat units"
XX
XX 98. .103
XX /label= synthetic linker
XX
XX 104
XX /label= artifact
XX /note= "see comments"
XX
XX 105. .209
XX /label= AAs 19-123 of CS protein
XX /note= "Region 1 contg. flanking region less signal
XX
XX 210
XX /label= artifact
XX /note= "see comments"
XX
XX 211. .335
XX /label= AAs 288-412 of CS protein
XX /note= "Region II flanking region"
XX
XX EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-00313257.
XX
XX 08-DEC-1989; 89US-00447746.
XX
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROS/) GROSS M S.
XX
XX Gross MS, Gordon DM, Hollingdal MR;
XX WPI; 1991-179771/25.
XX
XX Polypeptide comprising immunogenic determinants from P falciparum - for
```

PT vaccine against malaria infection in humans.
 PS Example 5; Page 11; 18pp; English.
 XX

CC The polypeptide is prep. by genetic engineering of genes encoding the P.
 CC falciparum circumsporozoite (CS) protein [Dane et al., Science 225 : 593
 CC (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et
 CC al., Nucleic Acids Research, 8 : 5845 (1980)]. The DNA encoding the 1st
 CC 81 AAs of the N-terminal of NS1 (NS1₈₁) is linked to a synthetic
 CC sequence encoding four repeat units (the variant form) from the
 CC immunodominant region, which in turn is linked via a synthetic sequence
 CC to DNA encoding Region I contg. flanking region less the 18 AA signal
 CC region. This is linked to DNA encoding Region II contg. flanking region.
 CC The Pro residue separating the Asp (at the C-terminal of the linker) from
 CC the Region I-contg. CS flanking region is an artifact of a filled-in
 CC BamHI site; the Gly separating the Region I and II-contg. CS flanking
 CC regions is an artifact of a synthetic FokI/NotI linker. The peptide
 CC can be used in a vaccine for protection against malaria. See also
 CC AAR1206-R12311 and AAR1375-R13178. (Updated on 25-MAR-2003 to correct
 CC PA field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX Sequence 335 AA;
 SQ

Query Match 49.7%; Score 91; DB 2; Length 335;
 Best Local Similarity 55.3%; Pred. No. 0.00013;
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 XX

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNS 36
 : : : : :
 DB 284 IKPSANKPKDELDYNDIEKKICKMEKCSSVFNVNS 321

RESULT 11
 ABO23530
 ID ABO23530 standard; protein; 396 AA.
 AC ABO23530;
 XX
 DT 04-SEP-2003 (first entry)
 XX
 DE Plasmodium falciparum outlier protein #2.
 XX
 KW Candidate protein identification; pathogen; anti-infective;
 KW outlier protein; virulence protein; antigen; drug target protein;
 KW pathogenic organism; antimicrobial.
 XX
 OS Plasmodium falciparum.
 XX
 PN US2003039963-A1.
 XX
 PD 27-FEB-2003.
 XX
 PF 30-MAR-2001; 2001US-00820843.
 XX
 PR 30-MAR-2001; 2001US-00820843.
 XX
 PA (BRAH/) BRAHMACHARI S K.
 PA (RAMA/) RAMACHANDRAN S.
 PA (NAND/) NANDI T.
 PA (BHIM/) BHIMARAO C.
 XX
 PI Brahmachari SK, Ramachandran S, Nandi T, Bhimarao C;
 XX
 DR WPI; 2003-492159/46.
 XX
 PT Identifying candidate proteins useful as anti-infectives involves
 PT matching outlier protein sequences with protein sequences in databases.
 XX
 PS Example 7; Page 44-45; 117pp; English.
 XX

CC The present invention relates to a method for identifying candidate
 CC proteins in pathogens useful as anti-infectives. The invention discloses
 CC a computational method which involves the calculation of several sequence

CC attributes and their subsequence analysis results in the identification
 CC of outlier proteins in different pathogens. The method is useful for the
 CC identification of outlier proteins (e.g. virulence proteins, antigens or
 CC proteins used as drug targets) in pathogenic organisms. The method of the
 CC invention provides reproducible results as it does not depend on the
 CC variable biochemical characterisation of proteins. ABO23500-ABO23617
 CC represent outlier proteins identified from different pathogenic organisms
 XX Sequence 396 AA;
 SQ

Query Match 49.7%; Score 91; DB 7; Length 396;
 Best Local Similarity 55.3%; Pred. No. 0.00015;
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 XX

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNS 36
 : : : : :
 DB 345 IKPSANKPKDELDYNDIEKKICKMEKCSSVFNVNS 382

RESULT 12
 AAP83144
 ID AAP83144 standard; protein; 411 AA.
 AC AAP83144;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-NOV-1990 (first entry)
 XX
 DE Sequence encoded by the circumsporozoite (CS) gene from Plasmodium
 DE falciparum.
 XX
 KW Vaccine; antigen; immunogen; probe; hybridisation; immunoassay;
 KW diagnosis.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Region 106..120
 FT Region /note="Region 1"
 FT Region 123..146
 FT Region /note="Repeat region, repeat unit = NANPNVDP"
 FT Region 147..206
 FT Region /note="Repeat region, repeat unit = NANP"
 FT Region 211..286
 FT Region /note="Repeat region, repeat unit = NANP"
 XX
 PN EP278940-A.
 XX
 PD 17-AUG-1988.
 XX
 PF 25-JAN-1988; 88EP-00870008.
 XX
 PR 30-JAN-1987; 87US-00009325.
 XX
 PA (SMIK) SMITH KLINE-RIT.
 PA (SKPK) SMITH KLINE-RIT.
 XX
 PI Cabazon T, De Wilde M, Harford N;
 XX
 DR WPI; 1988-229751/33.
 DR N-PSDB; AAN81108.
 XX
 PT DNA encoding Hepatitis B virus antigens and hybrids contg. them - used
 PT for expression in yeast to obtain vaccines and bivalent vaccines.
 XX
 PS Example; Fig 3Aa-3Af; 101pp; English.
 XX

CC Sequence of the CS gene (AAN81108) is from lambda-mpfl.. A recombinant DNA
 CC molecule is claimed, comprising functional DNA coding sequence fused, in
 CC phase, to a portion of the Pre-S2 region of a hepatitis B virus (HBV) Pre
 CC -S2-S protein coding sequence. The functional DNA coding sequence
 CC comprises the Pre-S2 coding sequence, Pre-S1 coding sequence or entire
 CC Pre-S1-Pre-S2 coding sequence, the CS protein coding sequence of

Best Local Similarity 55.3%; Pred. No. 0.00017;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
OY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVVNS 36
: : : : :
DB 152 IKPGSANKPKDLDYNDIEKKICKMEKCKSSVFNVVNS 189
: : : : :
RESULT 15
AAR37796
ID AAR37796 standard; protein; 424 AA.
XX AC AAR37796;
XX DT 25-MAR-2003 (revised)
XX DT 27-SEP-1993 (first entry)
XX DE RTS protein.
XX KW RTS, expression cassette; hybrid protein; S. cerevisiae; TDH3; cloning;
XX KW circumsporozoite protein; CSP; Plasmodium falciparum; strain 7G8;
XX KW hepatitis B virus; HBV; adv serotype; pres2 protein; S protein.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 1
XX FT /note= "Derived from S. cerevisiae TDH3 gene sequence"
XX FT Region 2..4
XX FT /note= "Cloning artefact"
XX FT Protein 5..193
XX FT /note= "Represents amino acids 210-398 of the CSP of P.
XX FT falciparum"
XX FT Region 194..197
XX FT /note= "Carboxy terminal amino acids from HBV (adv
XX FT serotype) pres2 protein"
XX FT Protein 198..424
XX FT /note= "S protein of HBV (adv serotype)"
XX PN W09310152-A1.
XX PD 27-MAY-1993.
XX PF 11-NOV-1992; 92WO-EP002591.
XX PR 16-NOV-1991; 91GB-00024390.
XX PR 27-FEB-1992; 92US-00842694.
XX PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI De Wilde M, Cohen J;
XX DR WPI; 1993-182494/22.
XX DR N-PSDB; AAQ42566.
XX PT Hybrid protein comprising plasmodium circumsporozoite protein and HBsAg -
XX PT useful as a vaccine for treating patients susceptible to Plasmodium
XX PT infections.
XX PS Disclosure; Fig 5; 59pp; English.
XX CC This sequence represents the RTS hybrid protein which is encoded by the
XX CC RTS expression cassette. This hybrid consists of a methionine residue
XX CC derived from S. cerevisiae TDH3 gene sequence, three amino acids, Met-Ala
XX CC -Pro, derived from a nucleotide sequence created by the cloning procedure
XX CC used to construct the hybrid gene, a stretch of 189 amino acids
XX CC representing amino acids 210 to 398 of the circumsporozoite protein (CSP)
XX CC of Plasmodium falciparum strain 7G8, an amino acid Arg created by the
XX CC cloning procedure, four amino acids, Pro-Val-Thr-Asn, representing the
XX CC four carboxy terminal residues of hepatitis B virus (HBV), adv serotype,
XX CC pres2 protein, and a stretch of 226 amino acids specifying the S protein
XX CC of HBV, adv serotype. This protein, and RTS* (see also AAR37797), may be
XX CC combined with an adjuvant and used in a vaccine for preventing plasmodium

CC infections. The vaccines produce a humoral response and also a cellular
CC immune response. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 424 AA;
Query Match 49.7%; Score 91; DB 2; Length 424;
Best Local Similarity 55.3%; Pred. No. 0.00017;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
OY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVVNS 36
: : : : :
DB 156 IKPGSANKPKDLDYNDIEKKICKMEKCKSSVFNVVNS 193
: : : : :
Search completed: March 10, 2004, 09:12:12
Job time : 54.2101 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 28.2957 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSSPSLDEKKIAMEKASSVFNVNS 36

Scoring table:

BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	182	99.5	36	9 US-09-848-834A-16	Sequence 15, Appl
2	182	99.5	51	9 US-09-848-834A-20	Sequence 20, Appl
3	95	51.9	20	9 US-09-848-834A-3	Sequence 3, Appl
4	95	51.9	33	9 US-09-848-834A-12	Sequence 12, Appl
5	91	43.7	396	10 US-09-820-843A-31	Sequence 31, Appl
6	89	48.6	19	14 US-10-239-313A-54	Sequence 54, Appl
7	89	48.6	21	10 US-09-932-165-1482	Sequence 1482, Ap
8	89	48.6	21	10 US-09-935-384-710	Sequence 710, App
9	89	48.6	21	10 US-09-942-052-711	Sequence 711, App
10	89	48.6	21	14 US-10-001-469-1404	Sequence 1404, Ap
11	89	48.6	21	14 US-10-128-711-97	Sequence 97, Appl
12	89	48.6	21	14 US-10-116-118-33	Sequence 33, Appl
13	89	48.6	21	14 US-10-082-109A-761	Sequence 761, App
14	89	48.6	21	14 US-10-005-480A-761	Sequence 761, App
15	89	48.6	21	14 US-10-277-292-652	Sequence 652, App

16 89 48.6 21 15 US-10-291-241-26 Sequence 26, Appl
17 89 48.6 21 15 US-10-280-340-652 Sequence 652, Appl
18 89 48.6 21 15 US-10-099-460-20 Sequence 20, Appl
19 89 48.6 21 15 US-10-024-652-2583 Sequence 2583, Ap
20 89 48.6 21 15 US-10-120-885A-25 Sequence 25, Appl
21 89 48.6 21 15 US-10-107-532-2 Sequence 2, Appl
22 89 48.6 21 15 US-10-121-016-52 Sequence 52, Appl
23 89 48.6 21 15 US-10-114-669-2 Sequence 2, Appl
24 89 48.6 21 15 US-10-120-835-44 Sequence 44, Appl
25 89 48.6 21 15 US-10-149-138-4224 Sequence 4224, Ap
26 89 48.6 21 15 US-10-114-432-39 Sequence 39, Appl
27 89 48.6 31 9 US-09-848-834A-15 Sequence 15, Appl
28 89 48.6 46 9 US-09-848-834A-19 Sequence 19, Appl
29 87 47.5 34 9 US-09-848-834A-13 Sequence 13, Appl
30 87 47.5 37 9 US-09-848-834A-14 Sequence 14, Appl
31 87 47.5 47 9 US-09-848-834A-17 Sequence 17, Appl
32 87 47.5 50 9 US-09-848-834A-18 Sequence 18, Appl
33 82 44.8 21 10 US-09-747-802-23 Sequence 23, Appl
34 82 44.8 21 10 US-09-865-294-15 Sequence 15, Appl
35 82 44.8 21 15 US-10-411-544-17 Sequence 17, Appl
36 82 44.8 31 15 US-10-411-544-39 Sequence 39, Appl
37 79 43.2 17 14 US-10-128-711-96 Sequence 96, Appl
38 79 43.2 17 14 US-10-239-313A-53 Sequence 53, Appl
39 75 41.0 16 15 US-10-372-111-3 Sequence 3, Appl
40 75 41.0 17 14 US-10-239-313A-36 Sequence 36, Appl
41 75 41.0 17 14 US-10-239-313A-38 Sequence 38, Appl
42 75 41.0 17 14 US-10-239-313A-39 Sequence 39, Appl
43 75 41.0 17 14 US-10-239-313A-41 Sequence 41, Appl
44 75 41.0 17 14 US-10-239-313A-60 Sequence 60, Appl
45 75 41.0 17 14 US-10-239-313A-64 Sequence 64, Appl

ALIGNMENTS

RESULT 1

US-09-848-834A-16
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Ashten Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
; OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(11)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; OTHER INFORMATION: (CSP) protein
US-09-848-834A-16

```
Query Match          99.5%; Score 182; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 8.8e-19;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVNS 36
    |||
DB 2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVNS 36

RESULT 2
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 20
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmodium
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          99.5%; Score 182; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVNS 36
    |||
DB 2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVNS 36

RESULT 3
US-09-848-834A-3
; Sequence 3, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
```

```
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium
; OTHER INFORMATION: falciparum circumsporozoite (CSP) protein
US-09-848-834A-3

Query Match          51.9%; Score 95; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFNVNS 36
    |||
DB 1 DEKKIAKMEKASSVFNVNS 20

RESULT 4
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of th
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD_RES
; LOCATION: (33)..(33)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; OTHER INFORMATION: (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(24)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (25)..(33)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-12

Query Match          51.9%; Score 95; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFNVNS 36
    |||
DB 1 DEKKIAKMEKASSVFNVNS 20
```

```
RESULT 5
; Sequence 31, Application US/09820843A
; Publication No. US2003039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; NAME/KEY: misc feature
; OTHER INFORMATION: Circumsporozoite (CS) protein
; NAME/KEY: misc feature
; OTHER INFORMATION: GI|4493889
US-09-820-843A-31

Query Match      49.7%; Score 91; DB 10; Length 396;
Best Local Similarity 55.3%; Pred. No. 0.00015;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGPSLD-----EKKIAKMEKASSVFNVNS 36
Db 345 IKPGSANKPKDELDYANDIEKKICKMEKSSVFNVNS 382

RESULT 6
US-10-239-313A-54
; Sequence 54, Application US/10239313A
; Publication No. US20030175285A1
; GENERAL INFORMATION:
; APPLICANT: KLINGUER - HAMOUR, Christine
; APPLICANT: CORVAIA, Nathalie
; APPLICANT: BECK, Alain
; APPLICANT: GOTSCH, Liliane
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS
; TITLE OF INVENTION: N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM
; TITLE OF INVENTION: OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID
; FILE REFERENCE: 343 727 - US
; CURRENT APPLICATION NUMBER: US/10/239,313A
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: FR 00/03711
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT 01/70772
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Plasmodium malariae
US-10-239-313A-54

Query Match      48.6%; Score 89; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
Db 1 EKKIAKMEKASSVFNVNS 19

RESULT 7
US-09-932-165-1482
; Sequence 1482, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: PARIS, MARY
; APPLICANT: SAFERAN, DOUGLAS
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AVA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CatF2E11 USEFUL IN TREATMENT AND
; TITLE OF INVENTION: DETECTION OF CANCER
; FILE REFERENCE: 51158-20033.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1482
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-932-165-1482

Query Match      48.6%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 8
US-09-935-384-710
; Sequence 710, Application US/09935384
; Publication No. US20030166526A1
; GENERAL INFORMATION:
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: PARIS, MARY
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1H4
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20033.00
; CURRENT APPLICATION NUMBER: US/09/935,384
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 783
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 710
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-935-384-710

Query Match      48.6%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
Db 3 EKKIAKMEKASSVFNVNS 21
```


RESULT 9
US-09-942-052-711
; Sequence 711, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Arar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 711
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-942-052-711

Query Match 48.6%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 10
US-10-001-469-1404
; Sequence 1404, Application US/10001469
; Publication No. US20030091562A1
; GENERAL INFORMATION:
; APPLICANT: JAKOBOVITS, AYA
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: SAFFRAN, DOUGLAS
; APPLICANT: HUBERT, RENE
; APPLICANT: FARIS, MARY
; APPLICANT: CHALLITA-EID, PIA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED
; TITLE OF INVENTION: 101P3A11 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20024.20
; CURRENT APPLICATION NUMBER: US/10/001,469
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/291,118
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 09/680,728
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2888
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1404
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-001-469-1404

Query Match 48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36

DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 11
US-10-128-711-97
; Sequence 97, Application US/10128711
; Publication No. US20030099634A1
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/10/128,711
; FILING DATE: 22-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; 378-398"
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-10-128-711-97

Query Match 48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
DB 3 EKKIAKMEKASSVFNVNS 21

```

RESULT 12
US-10-116-118-33
; Sequence 33, Application US/10116118
; Publication No. US20030143672A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakovovits, Aya
; TITLE OF INVENTION: Heteroclitic Analogs and Related Methods
; FILE REFERENCE: 2060 0090003
; CURRENT APPLICATION NUMBER: US/10/116,118
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/166,529
; PRIOR FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: US 60/239,008
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-116-118-33

Query Match      48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 EKKIARMEKASSVFNVNS 36
Db      3 EKKIARMEKASSVFNVNS 21

RESULT 13
US-10-062-109A-761
; Sequence 761, Application US/10062109A
; Publication No. US20030165505A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakovovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20062.01
; CURRENT APPLICATION NUMBER: US/10/062,109A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 761
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-109A-761

Query Match      48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 EKKIARMEKASSVFNVNS 36
Db      3 EKKIARMEKASSVFNVNS 21

RESULT 14
US-10-005-480A-761
; Sequence 761, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakovovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20062.00
; CURRENT APPLICATION NUMBER: US/10/005,480A
; CURRENT FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 761
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-005-480A-761

Query Match      48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 EKKIARMEKASSVFNVNS 36
Db      3 EKKIARMEKASSVFNVNS 21

RESULT 15
US-10-277-292-652
; Sequence 652, Application US/10277292
; Publication No. US20030199470A1
; GENERAL INFORMATION:
; APPLICANT: FARRIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APEAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOVITZ, AYA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PLD7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/277,292
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 652
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-277-292-652

Query Match      48.6%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 EKKIARMEKASSVFNVNS 36
Db      3 EKKIARMEKASSVFNVNS 21

```

Wed Mar 10 10:34:22 2004

us-09-848-834a-16.open.rapb

Page 6

Db 3 EKKIAKMEKASSVFNVNS 21

Search completed: March 10, 2004, 10:25:49
Job time : 28.2957 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 12.0019 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWGYGLRPGSSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	84	51.9	24	4	US-08-464-496-31		Sequence 31, Appl
2	84	51.9	24	5	PCT-US92-07218-31		Sequence 31, Appl
3	78	48.1	182	4	US-09-396-937-16		Sequence 16, Appl
4	77	47.5	173	4	US-09-396-937-18		Sequence 18, Appl
5	76	46.9	17	4	US-08-472-701-23		Sequence 23, Appl
6	76	46.9	17	5	PCT-US95-08596-23		Sequence 23, Appl
7	76	46.9	31	5	PCT-US93-11703-63		Sequence 63, Appl
8	76	46.9	37	1	US-08-446-692-57		Sequence 57, Appl
9	76	46.9	37	1	US-08-446-692-63		Sequence 63, Appl
10	76	46.9	37	2	US-08-488-351A-57		Sequence 57, Appl
11	76	46.9	37	2	US-08-488-351A-63		Sequence 63, Appl
12	76	46.9	47	1	US-08-446-692-35		Sequence 35, Appl
13	76	46.9	47	2	US-08-488-351A-35		Sequence 35, Appl
14	76	46.9	853	4	US-08-913-880C-17		Sequence 17, Appl
15	76	46.9	858	4	US-08-913-880C-15		Sequence 15, Appl
16	76	46.9	860	4	US-08-913-880C-16		Sequence 16, Appl
17	76	46.9	862	4	US-08-913-880C-14		Sequence 14, Appl
18	76	46.9	865	4	US-08-913-880C-13		Sequence 13, Appl
19	76	46.9	866	4	US-08-913-880C-12		Sequence 12, Appl
20	76	46.9	874	4	US-08-913-880C-11		Sequence 11, Appl
21	76	46.9	875	4	US-08-913-880C-10		Sequence 10, Appl
22	76	46.9	1315	4	US-08-913-880C-10		Sequence 10, Appl
23	74	45.7	15	2	US-08-319-704-10		Sequence 10, Appl
24	74	45.7	15	2	US-08-661-052-6		Sequence 6, Appl
25	74	45.7	15	2	US-08-460-502-7		Sequence 7, Appl
26	74	45.7	15	3	US-09-046-373-2		Sequence 2, Appl
27	74	45.7	15	3	US-09-188-082-6		Sequence 6, Appl

28 74 45.7 15 4 US-09-364-088-6 Sequence 6, Appl
29 74 45.7 15 4 US-09-102-716-6 Sequence 6, Appl
30 74 45.7 15 4 US-09-148-711A-7 Sequence 7, Appl
31 74 45.7 15 4 US-08-716-249-4 Sequence 4, Appl
32 74 45.7 15 4 US-09-362-731A-13 Sequence 13, Appl
33 74 45.7 15 4 US-09-396-937-34 Sequence 34, Appl
34 74 45.7 15 4 US-09-405-986A-1 Sequence 1, Appl
35 74 45.7 15 5 PCT-US93-11703-69 Sequence 69, Appl
36 74 45.7 16 3 US-09-248-588-55 Sequence 55, Appl
37 74 45.7 17 1 US-08-446-692-4 Sequence 4, Appl
38 74 45.7 17 2 US-08-488-351A-4 Sequence 4, Appl
39 74 45.7 17 3 US-09-100-409A-40 Sequence 40, Appl
40 74 45.7 17 5 PCT-US95-13841-7 Sequence 7, Appl
41 74 45.7 27 1 US-08-446-692-13 Sequence 13, Appl
42 74 45.7 27 2 US-08-488-351A-13 Sequence 13, Appl
43 74 45.7 31 4 US-09-362-731A-1 Sequence 1, Appl
44 74 45.7 32 4 US-09-362-731A-5 Sequence 5, Appl
45 74 45.7 50 4 US-08-945-289-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-464-496-31
; Sequence 31, Application US/08464496
; Patent No. 6322789
; GENERAL INFORMATION:
; APPLICANT: Epimmune, Inc.
; APPLICANT: Vitiello, Maria
; APPLICANT: Chesnut, Robert
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; FILE REFERENCE: 39963-20001.13
; CURRENT APPLICATION NUMBER: US/08/464,496
; CURRENT FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 07/935,811
; PRIOR FILING DATE: 1992-08-26
; PRIOR APPLICATION NUMBER: 07/874,491
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/827,682
; PRIOR FILING DATE: 1992-01-29
; PRIOR APPLICATION NUMBER: 07/749,568
; PRIOR FILING DATE: 1991-08-26
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligopeptide
US-08-464-496-31

Query Match 51.9%; Score 84; DB 4; Length 24;
Best Local Similarity 77.3%; Pred. No. 6.6e-06;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 PGSSGPSLQYIKANSKFIGITE 30
Db 3 PSDFFSVQYIKANSKFIGITE 24

RESULT 2
PCT-US92-07218-31
; Sequence 31, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-092DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-472-701-23

Query Match 46.9%; Score 76; DB 4; Length 17;
Best Local Similarity 93.8%; Pred. No. 7e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31
Db 2 MQYIKANSKFIGITEL 17

RESULT 6
PCT-US95-08596-23
Sequence 23, Application PC/TUS9508596
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Proinsulin Peptide Compounds for Detecting
and Treating Type I Diabetes
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08596
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/272,220
FILING DATE: 08-JULY-1994

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-092PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-08596-23

Query Match 46.9%; Score 76; DB 5; Length 17;
Best Local Similarity 93.8%; Pred. No. 7e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31
Db 2 MQYIKANSKFIGITEL 17

RESULT 7
PCT-US93-11703-63
Sequence 63, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-63

Query Match 46.9%; Score 76; DB 5; Length 31;
Best Local Similarity 93.8%; Pred. No. 0.00014;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31
Db 9 MQYIKANSKFIGITEL 24

RESULT 8

US-08-446-692-57
Sequence 57, Application US/08446692
Patent No. 5759551

GENERAL INFORMATION:

APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY

```

; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-57

Query Match 46.9%; Score 76; DB 1; Length 37;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLQYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 9
US-08-446-692-63
; Sequence 63, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid

QY 13 GPSLQYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 10
US-08-488-351A-57
; Sequence 57, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-57

Query Match 46.9%; Score 76; DB 2; Length 37;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLQYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35
```

Db 17 GKKQYIKANSKFIGITEL 35

RESULT 11

US-08-488-351A-63
; Sequence 63, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-63

Query Match 46.9%; Score 76; DB 2; Length 37;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GPSLOYIKANSKFIGITEL 31

Db 1 GKKQYIKANSKFIGITEL 19

RESULT 12

US-08-446-692-35
; Sequence 35, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-35

Query Match 46.9%; Score 76; DB 1; Length 47;
Best Local Similarity 84.2%; Pred. No. 0.00022;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GPSLOYIKANSKFIGITEL 31

Db 17 GKKQYIKANSKFIGITEL 35

RESULT 13

US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692


```
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35
```

```
Query Match 46.9%; Score 76; DB 2; Length 47;
Best Local Similarity 84.2%; Pred. No. 0.0022;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 13 GPSLOYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35
```

RESULT 14

```
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
; US-08-913-880C-17
```

```
Query Match 46.9%; Score 76; DB 4; Length 853;
Best Local Similarity 93.8%; Pred. No. 0.0061;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LOYIKANSKFIGITEL 31
Db 367 MQYIKANSKFIGITEL 382
```

RESULT 15

```
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
```

```
; CURRENT FILING DATE: 1997-11-24
```

```
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 16
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
; US-08-913-880C-16
```

```
Query Match 46.9%; Score 76; DB 4; Length 858;
Best Local Similarity 93.8%; Pred. No. 0.0061;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LOYIKANSKFIGITEL 31
Db 372 MQYIKANSKFIGITEL 387
```

```
Search completed: March 10, 2004, 09:28:55
Job time : 12.0019 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 5.66926 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRPGSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	46.9	1314	1	TETX CLOTE
2	61	37.7	63	1	GON1_MESAU
3	60	37.0	61	1	GON1_SHEEP
4	59.5	36.7	90	1	GON1_RANCA
5	58	35.8	67	1	GON1_MACMU
6	58	35.8	89	1	GON1_XENIA
7	58	35.8	90	1	GON1_MOUSE
8	58	35.8	91	1	GON1_PIG
9	58	35.8	92	1	GON1_HUMAN
10	58	35.8	92	1	GON1_RAT
11	58	35.8	92	1	GON1_TURGE
12	55.5	34.3	66	1	VGB8_BPML5
13	54	33.3	10	1	GON1_ALIMI
14	54	33.3	92	1	GON1_CHICK
15	52	32.1	94	1	GON1_HAPBU
16	52	32.1	95	1	GON1_MORSA
17	52	32.1	95	1	GON1_PAGMA
18	52	32.1	95	1	GON1_SPAU
19	52	32.1	99	1	GON1_DICLA
20	52	32.1	99	1	ACEA_LYCES
21	50.5	31.2	575	1	GON3_DICLA
22	50	30.9	80	1	GON1_CLAGA
23	50	30.9	90	1	GON8_RANDY
24	50	30.9	249	1	FRA_MYCLE
25	49	30.2	29	1	GON1_CAVPO
26	49	30.2	190	1	HISE_PYRAE
27	49	30.2	293	1	KHSE_PYRHO
28	49	30.2	294	1	KHSE_PYRAB
29	49	30.2	408	1	SEPR_THESR
30	49	30.2	444	1	GARP_ECOLI
31	49	30.2	485	1	RT16_MYXXA
32	48.5	29.9	892	1	1F2_YERPE
33	48	29.6	91	1	GON1_ORYLA

34 48 29.6 390 1 YB12_SCHPO
35 47.5 29.3 89 1 GON3_PORNO
36 47.5 29.3 90 1 GON3_HAPBU
37 47.5 29.3 90 1 GON3_ORYLA
38 47.5 29.3 90 1 GON3_PAGMA
39 47.5 29.3 90 1 GON3_SPAU
40 47.5 29.3 276 1 SFAS_CHLRE
41 47.5 29.3 407 1 VG10_HSVSA
42 47.5 29.3 526 1 MALQ_CHLPN
43 47 29.0 10 1 GON1_CLUPA
44 47 29.0 120 1 HV03_MOUSE
45 47 29.0 140 1 HV02_MOUSE

P87167 schizosacch
P51922 porichthys
P45652 haplochromi
Q9dd49 o progonado
P51921 pagrus majo
P51923 sparus auro
Q39618 chlamydomon
P24913 herpesvirus
Q92812 chlamydia p
P81749 clupea pall
P01747 mus musculu
P01746 mus musculu

ALIGNMENTS

RESULT 1
TETX CLOTE
ID TETX CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxilyein) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
OS TETX OR CTP60.
GN Clostridium tetani.
OG Plasmid p88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp.
RX MEDLINE=87053514; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.,
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=pE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RN [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
J. Bacteriol. 165:21-27(1986).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulphydryl groups
RT in tetanus toxin.";

RL Eur. J. Biochem. 188:39-45(1990).
 RN [6]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
 RT "limited proteolysis of tetanus toxin. Relation to activity and
 RT identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51(1991).
 RN [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "tetanus toxin is a zinc protein and its inhibition of
 RT neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583(1992).
 RN [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.P., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 RT neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOSOMAL-2.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 CC synaptobrevin 2.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -1- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X04436; CAA28033.1; -;
 CC EMBL; X06214; CAA29564.1; -;
 CC EMBL; AF528097; AAC37454.1; -;
 CC EMBL; M12739; AAA23282.1; -;
 CC PIR; A25689; BTCLTN.
 CC PDB; 1AF9; 29-APR-98.
 CC PDB; 1A8D; 14-OCT-98.
 CC PDB; 1DOH; 27-MAR-00.
 CC PDB; 1DFQ; 24-MAR-00.
 CC PDB; 1DIW; 24-MAR-00.
 CC PDB; 1DLL; 24-MAR-00.
 CC PDB; 1FV3; 05-SEP-01.
 CC MEROPS; M27.001; -;
 CC InterPro; IPR006985; ConA like lec_g1.
 CC InterPro; IPR002160; Kunitz_leguene.
 CC InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR000395; Peptidase M27.
 DR Pfam; PF01742; Peptidase M27; 1.
 DR PRINTS; PR00760; BONTOKILYSIN.
 DR ProDom; PD001963; Bontokilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT_MET 0 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT METAL 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT ACT_SITE 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 233 233 BY SIMILARITY.
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 226 246 POTENTIAL.
 FT TRANSMEM 669 689 POTENTIAL.
 FT DISULFID 438 466 INTERCHAIN.
 FT DISULFID 1076 1092
 FT HELIX 876 882
 FT TURN 883 883
 FT STRAND 884 891
 FT TURN 892 893
 FT STRAND 894 897
 FT STRAND 904 907
 FT TURN 909 910
 FT STRAND 912 915
 FT STRAND 920 925
 FT TURN 928 929
 FT STRAND 932 935
 FT HELIX 938 940
 FT TURN 941 946
 FT STRAND 949 956
 FT HELIX 962 968
 FT TURN 969 970
 FT STRAND 972 977
 FT STRAND 980 981
 FT HELIX 983 985
 FT STRAND 987 995
 FT TURN 996 997
 FT STRAND 998 1004
 FT TURN 1006 1007
 FT STRAND 1010 1016
 FT STRAND 1020 1020
 FT TURN 1021 1022
 FT STRAND 1031 1037
 FT TURN 1039 1040
 FT STRAND 1042 1047
 FT TURN 1048 1049
 FT STRAND 1050 1056
 FT TURN 1058 1059
 FT STRAND 1068 1074
 FT TURN 1079 1080
 FT STRAND 1082 1091
 FT HELIX 1097 1105
 FT TURN 1106 1107
 FT STRAND 1112 1112
 FT STRAND 1114 1114
 FT TURN 1116 1117
 FT STRAND 1120 1120
 FT STRAND 1122 1122
 FT TURN 1123 1124
 FT STRAND 1127 1131
 FT HELIX 1132 1134
 FT TURN 1135 1136
 FT STRAND 1137 1141
 FT TURN 1144 1145
 FT STRAND 1148 1152
 FT STRAND 1155 1158
 FT TURN 1159 1162
 FT STRAND 1163 1166
 FT STRAND 1173 1178
 FT TURN 1184 1185
 FT STRAND 1188 1188
 FT STRAND 1190 1190

Query Match 46.9%; Score 76; DB 1; Length 1314;
 Best Local Similarity 93.8%; Pred. No. 0.0045;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LQYIKANSKFIGITEL 31
 :|||||
 DB 828 MQYIKANSKFIGITEL 843

RESULT 2
 GON1 MESAU STANDARD; PRT; 63 AA.
 AC O09163;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
 DE (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Mesocricetus.
 OC NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
 RA Submitted (MAR-1997) to the EMBL/GenBank/DBSJ databases.
 RL -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U91938; AB51302.1; .
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadolibirinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBIRNI.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1
 FT CHAIN 1 >63
 FT PEPTIDE 1 10
 FT ACT_SITE 14 >63
 FT PROGNADOLIBIRIN I.
 FT GONADOLIBIRIN I.
 FT GNRH-ASSOCIATED PEPTIDE I (BY
 FT SIMILARITY).
 FT ACT_SITE 3 3
 FT APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY (BY SIMILARITY).
 FT MOD_RES 1 1
 FT PYRROLIDONE CARBOXYLIC ACID (BY
 FT SIMILARITY).
 FT MOD_RES 10 10
 FT AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
 FT SIMILARITY).
 FT NON_TER 63 63
 FT SEQUENCE 63 AA; 7370 MW; FC9499567677180 CRC64;

Query Match 37.7%; Score 61; DB 1; Length 63;
 Best Local Similarity 46.2%; Pred. No. 0.023;
 Matches 12; Conservative 3; Mismatches 3; Indels 8; Gaps 1;

QY 2 HWSYGLRPGSSGPSLQYIKANSKFIG 27
 :|||||

DB 2 HWSYGLRPGG-----KRNAERLG 19

RESULT 3

GON1 SHEEP STANDARD; PRT; 61 AA.
 AC Q28588;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
 DE (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE OF 12-61 FROM N.A.
 RC STRAIN=Western range; TISSUE=Hypothalamus;
 RA Rodriguez R.E., Wise M.B.;
 RA Submitted (OCT-1993) to the EMBL/GenBank/DBSJ databases.
 RN [2]
 RP SEQUENCE OF 1-10.
 RX MEDLINE=72094314; PubMed=4550508;
 RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
 RA Fellows R., Blackwell R., Vale W., Guillemin R.;
 RT "Primary structure of the ovine hypothalamic luteinizing hormone-
 RT releasing factor (LRH) (LH-hypothalamus-LRF-gas chromatography-mass
 RT spectrometry-decapeptide-Edman degradation).";
 RT Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U02517; AAA03433.1; .
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadolibirinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBIRNI.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1
 FT CHAIN 1 >61
 FT PEPTIDE 1 10
 FT GONADOLIBIRIN I.
 FT GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 14 >61
 FT APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 1 1
 FT PYRROLIDONE CARBOXYLIC ACID.
 FT AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61 61
 FT SEQUENCE 61 AA; 6828 MW; 53962A1AE319B9F0 CRC64;

Query Match 37.0%; Score 60; DB 1; Length 61;
 Best Local Similarity 52.2%; Pred. No. 0.032;
 Matches 12; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 2 HWSYGLRPGSSGPSLQYIKANSK 24
 :|||||

Db 2 HWSYGLRPGG-----KRNAX 16

```
RESULT 4
GONI_RANCA
ID GONI_RANCA STANDARD; PRT; 90 AA.
AC Q90763;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I (GAP1)].
GN GNRH1 OR GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OC NCBI_TaxID=8400;
OX [1]
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Forebrain; PubMed=11170016;
RX MEDLINE=21102951; Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Wang L., Yoo M.S., Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GNRH1 and GNRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201 (2001).
CC -|- FUNCTION: Stimulates the secretion of gonadotropins (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Forebrain.
CC -|- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC during post-breeding. Not expressed in pituitary.
CC -|- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF188754; AL05972.1; -.
CC DR GO:0005576; C:extracellular; NAS.
CC DR GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
CC DR GO:0007755; P:hormone mediated signaling; NAS.
CC DR GO:0000003; P:reproduction; NAS.
CC DR InterPro; IPR002012; GnRH.
CC DR Pfam; PF00446; GnRH; 1.
CC DR PRINTS; PR01541; GONADOLIBERNI.
CC DR PROSITE; PS00473; GNRH; 1.
CC DR Cleavage on pair of basic residues; Hormone; Amidation; Signal;
CC SIGNAL 1 24 POTENTIAL.
CC FT CHAIN 1 24 PROGNADOLIBERIN I.
CC FT PEPTIDE 25 34 GONADOLIBERIN I.
CC FT PEPTIDE 25 34 GNRH-ASSOCIATED PEPTIDE I (BY
CC FT PEPTIDE 38 86 SIMILARITY).
CC FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID (BY
CC FT MOD_RES 34 34 SIMILARITY).
CC FT MOD_RES 34 34 AMIDATION (G-35 PROVIDE AMIDE GROUP) (BY
CC FT SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;
CC
CC Query Match 36.7%; Score 59.5; DB 1; Length 90;
CC Best Local Similarity 47.1%; Pred. No. 0.06;
CC Matches 16; Conservative 4; Mismatches 9; Indels 5; Gaps 3;
CC
CC QY 2 HWSYGLRPGSG--PSLQ--YIKANSKFIGITEL 31
CC Db ||||| : : : : :
CC 26 HWSYGLRPGGKREVESLQESYAEVPNE-VSFTEL 58
CC ||||| : : : : :

```

```
RESULT 5
GONI_MACMU
ID GONI_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359 (1994).
CC -|- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S75918; AB33096.1; -.
CC DR PIR; I78541; I78541.
CC DR InterPro; IPR002012; GnRH.
CC DR InterPro; IPR004073; GonadoliberinI.
CC DR Pfam; PF00446; GnRH; 1.
CC DR PRINTS; PR01541; GONADOLIBERNI.
CC DR PROSITE; PS00473; GNRH; 1.
CC DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Pyrrolidone carboxylic acid.
CC FT NON_TER 1 1
CC FT SIGNAL <1 5 BY SIMILARITY.
CC FT CHAIN 6 >67 PROGNADOLIBERIN I.
CC FT PEPTIDE 6 15 GONADOLIBERIN I.
CC FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
CC FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC FT MOD_RES 6 6 ACTIVITY (BY SIMILARITY).
CC FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
CC FT MOD_RES 15 15 SIMILARITY).
CC FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
CC FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
CC
CC Query Match 35.8%; Score 58; DB 1; Length 67;
CC Best Local Similarity 100.0%; Pred. No. 0.071;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 HWSYGLRPG 10
CC Db |||||
CC 7 HWSYGLRPG 15

```

```

DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
DE GN GN RH1 OR GNRH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolic K., Seeburg P.H.;
RA "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse.";
RL Science 234:1366-1371(1986).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14872; AAA37717.1; -.
DR PIR; A47578; RHMSG.
DR MGD; MGI:95789; GnRH.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 21
FT CHAIN 22 90
FT PEPTIDE 22 31
FT PEPTIDE 35 90
FT ACT_SITE 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
FT SEQUENCE 90 AA; 10337 MW; 1C0766FA4826B4D9 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 23 HWSYGLRPG 31
RESULT 8
GONI_PIG
ID GONI_PIG STANDARD; PRT; 91 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GN RH1 OR GNRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```

RESULT 6
GONI_XENLA
ID GONI_XENLA STANDARD; PRT; 89 AA.
DT P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OX Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Forebrain;
RX MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battley J.F.;
RA "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GnRH-associated peptide, but brain onset is delayed until
RT metamorphosis.";
RL Endocrinology 134:1835-1844(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L28040; AAA49728.1; -.
DR PIR; I51423; I51423.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT PEPTIDE 37 89
FT MOD_RES 37 85
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 7
GONI_MOUSE
ID GONI_MOUSE STANDARD; PRT; 90 AA.
DT P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)

```

OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463 (1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method.";
 RL Biochem. Biophys. Res. Commun. 45:822-827 (1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487 (1971).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L32864; AAA31066.1; -.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 35.8%; Score 58; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 |||||
 DB 25 HWSYGLRPG 33
 RESULT 9
 GONI_HUMAN
 ID GONI_HUMAN STANDARD; PRT; 92 AA.
 AC P0148;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 DE Luteinizing hormone-releasing hormone I] (Gonadoliberin-releasing
 DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
 DE peptide II.
 GN GnRH OR GnRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene.";
 RL Nucleic Acids Res. 17:6403-6403 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183 (1986).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT SER-16.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone.";
 RL Nature 311:666-668 (1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Roussseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071 (1982).
 RN [5]
 RP VARIANT SER-16.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238 (1999).
 RN [6]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373 (1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelaf (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; X01059; CAA25526.1; -
DR EMBL; M12578; AAA35916.1; -
DR EMBL; X15215; CAA33285.1; -
DR PIR; S05308; RHUG.
DR Genew; HGNC:4419; GNRH1.
DR MIM; 152760; -
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GNRH.1.
DR PRINTS; PR01541; GONADOLIBERN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal; Polymorphism;
KW Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT VARIANT 16 16
FT SEQUENCE 92 AA; 10380 MW; 30A722221B076FA79 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 10
GONI_RAT ID GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (LH-RH I) (Luliberin I); Prolactin release-inhibiting factor I.
DE GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183 (1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476569;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic expression."
RL Mol. Endocrinol. 3:1257-1262 (1989).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Bialock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone-releasing hormone mRNA."
RL Cell. Mol. Neurobiol. 12:447-454 (1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same DNA locus."
RL Science 235:1514-1517 (1987).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Central nervous system.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements/ or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; S05870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RHRTG.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DABA3EB3 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 11
GONI_TUPGB ID GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (Luliberin I); Prolactin release-inhibiting factor I.

```


DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GnRH1 OR GnRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=37347;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasren T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Ferrel R.D.;
RT Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal.
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U63326; AAB16837.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin.
DR Pfam; PF00446; GnRH; 1.
DR PRINIS; PR01541; GONADOLIBERIN.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 24 33 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 12
VG84 BPML5 STANDARD; PRT; 66 AA.
AC Q05301;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Gene 84 protein (GP84).
GN 84.
OS Mycobacteriophage L5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC L5-like viruses
OX NCBI_TaxID=31757;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=932111282; PubMed=8459766;
RA Hatfull G.F., Sarkis G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
RT a phage system for mycobacterial genetics.";
RL Mol. Microbiol. 7:395-405(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z18946; CAA79460.1; -.
DR PIR; S31029; S31029.
SQ SEQUENCE 66 AA; 7424 MW; 9C7104C7A4FA74A5 CRC64;
Query Match 34.3%; Score 55.5; DB 1; Length 66;
Best Local Similarity 43.8%; Pred. No. 0.17;
Matches 14; Conservative 2; Mismatches 9; Indels 9; Gaps 2;
QY 5 YGL-----RPGSGPSLQYIKANSKFIGITEL 31
DB 36 YGFEVDWYEPGESG-----YIKRNGKFGVTWEV 63
RESULT 13
GONI_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylia; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RN SEQUENCE
RC TISSUE=Brain;
RX MEDLINE=9132338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
FT MOD_RES 1 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284E23D7286B45A3 CRC64;
Query Match 33.3%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.031;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10
RESULT 14
GONI_CHKCK

D	P37042; F20407;	STANDARD;	PRT;	92 AA.
AC	01-FEB-1991 (Rel. 17, Created)			
CT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
FT	Gonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) luteinizing hormone-releasing hormone I] (Gonadotropin-releasing hormone I) (GNRH I) (luliberin I); GNRH-associated peptide I].			
DE	Gallus gallus (Chicken).			
ES	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OS	Archoseauria; Aves; Neognathae; Galliformes; Phasianinae;			
OC	Gallus.			
NCBI_TaxID=9031;				
[1]	SEQUENCE FROM N.A.			
RN	STRAIN=White leghorn;			
RP	MEDLINE=94059355; PubMed=7903095;			
RX	Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;			
RA	"Characterization of the chicken pregonadotrophin-releasing hormone-I gene";			
RT	J. Mol. Endocrinol. 11:19-29(1993).			
RL	[2]			
RN	SEQUENCE OF 24-33.			
RP	TISSUE=Hypothalamus;			
RX	MEDLINE=82265778; PubMed=7050119;			
RA	King J.A., Millar R.P.;			
RT	"Structure of chicken hypothalamic luteinizing hormone-releasing hormone. II. Isolation and characterization.";			
RL	J. Biol. Chem. 257:10729-10732(1982).			
[3]				
RN	SEQUENCE OF 24-33.			
RP	TISSUE=Hypothalamus;			
RX	King J.A., Millar R.P.;			
RA	"Structure of avian hypothalamic gonadotrophin-releasing hormone.";			
RL	S. Afr. J. Sci. 78:124-125(1982).			
[4]				
RN	SYNTHESIS OF 24-33.			
RP	MEDLINE=82265777; PubMed=7050118;			
RX	King J.A., Millar R.P.;			
RA	"Structure of chicken hypothalamic luteinizing hormone-releasing hormone. I. Structural determination on partially purified material";			
RL	J. Biol. Chem. 257:10722-10728(1982).			
[5]				
RN	-!- FUNCTION: Stimulates the secretion of gonadotropins.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: Belongs to the GNRH family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X59491; CAA49246.1; ..			
DR	PIR; I50644; I50644.			
DR	GO; GO:0005178; Extracellular; IDA.			
DR	GO; GO:0005183; Fluteinizing hormone-releasing factor activity; IDA.			
DR	GO; GO:0007275; P:development; IDA.			
DR	InterPro; IPR002012; GNRH.			
DR	InterPro; IPR004079; GonadoliberinI.			
DR	Pfam; PF00446; GNRH; 1.			
DR	PRINTS; PR01541; GONADOLIBERNI.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
XW	Signal; Pyrolydine carboxylic acid.			
FT	SIGNAL 1 23			
FT	CHAIN 24 92			
FT	PEPTIDE 24 33			
FT	PEPTIDE 37 92			
FT	PEPTIDE 24 24			
FT	MOD_RES 24 33			
FT	MOD_RES 33			
FT	AMIDATION (G-34 PROVIDE AMIDE GROUP).			

SQ	SEQUENCE	92 AA;	10206 NW;	61AE7EBAF508B6A	CRC64;
Query Match	Similarity	33.3%;	Score 54;	DB 1;	Length 92;
Best Local Similarity	88.9%;	Pred. No. 0.41;			
Matches	8;	Conservative	1;	Mismatches	0;
				Indels	0;
				Gaps	0;
Ov	2	HWSYGLRPG	10		
		:			
Db	25	HWSYGLQPG	33		
		:			
RESULT 15					
GONI_HAPBU					
ID	GONI_HAPBU	STANDARD;	PRT;	94 AA.	
AC	P51918; O93387;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Gonadoliberin I precursor [Contains: Gonadoliberin I (Luteinizing hormone releasing hormone I) (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I) (luliberin I); GNRH-associated peptide I].				
DE	GNRH1.				
GN	Haplochromis burtoni (Burton's mouthbrooder).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Neoteleostei;				
OC	Acanthomorpha; Acanthopterygii; Perciformes; Labroides;				
OC	Cichlidae; Astacotiropia.				
OX	NCBI_TaxID=8153;				
[1]					
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=95396797; PubMed=7667296;				
RX	White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;				
RA	"Three gonadotropin-releasing hormone genes in one organism suggest novel roles for an ancient peptide.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).				
RL	[2]				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=99061842; PubMed=9843638;				

D	P37042; F20407;	STANDARD;	PRT;	92 AA.
AC	01-FEB-1991 (Rel. 17, Created)			
CT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
FT	Gonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) luteinizing hormone-releasing hormone I]			
DE	"Characterization of the chicken pregonadotrophin-releasing hormone-I gene";			
EE	J. Mol. Endocrinol. 11:19-29(1993).			
ES	[2]			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
CC	Gallus.			
NCBI_TaxID=9031;				
GN	[1]			
EN	SEQUENCE FROM N.A.			
RN	STRAIN=White leghorn;			
RP	MEDLINE=94059355; PubMed=7903095;			
RX	Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;			
RA	"Characterization of the chicken pregonadotrophin-releasing hormone-I gene";			
RT	J. Biol. Chem. 257:10729-10732(1982).			
RL	[3]			
RN	SEQUENCE OF 24-33.			
RP	TISSUE=Hypothalamus;			
RX	MEDLINE=82265777; PubMed=7050119;			
RA	King J.A., Millar R.P.;			
RT	"Structure of avian hypothalamic gonadotrophin-releasing hormone.";			
RL	S. Afr. J. Sci. 78:124-125(1982).			
RN	[4]			
RP	SYNTHESIS OF 24-33.			
RX	MEDLINE=82265777; PubMed=7050119;			
RA	King J.A., Millar R.P.;			
RT	"Structure of chicken hypothalamic luteinizing hormone-releasing hormone. I. Structural determination on partially purified material";			
RL	J. Biol. Chem. 257:10722-10728(1982).			
CC	-!- FUNCTION: Stimulates the secretion of gonadotropins.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: Belongs to the GnRH family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X59491; CAA49246.1; ..			
DR	PIR; I50644; I50644.			
DR	GO; GO:0005176; Extracellular; IDA.			
DR	GO; GO:0005183; Fluteinizing hormone-releasing factor activity; IDA.			
DR	GO; GO:0007275; P:development; IDA.			
DR	InterPro; IPR002012; GnRH.			
DR	InterPro; IPR004079; GonadoliberinI.			
DR	Pfam; PF00446; GnRH; 1.			
DR	PRINTS; PR01541; GONADOLIBERNI.			
DR	PROSITE; PS00473; GnRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
XW	Signal; Pyrolydine carboxylic acid.			
FT	SIGNAL 1 23			
FT	CHAIN 24 92			
FT	PEPTIDE 24 92			
FT	PEPTIDE 37 92			
FT	MOD_RES 24 24			
FT	RES 33 33			
FT	AMIDATION (G-34 PROVIDE AMIDE GROUP).			

Query Match 33.3%; Score 54; DB 1; Length 92;
 Best Local Similarity 88.9%; Pred. No. 0.41;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QV 2 HWSYGLRPG 10
 Db 25 HWSYGLQPG 33
 |||||:
 |||||:

DR PIR; I50739; I50739.
DR GO; GO:000576; C:extracellular; NAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
DR GO; GO:0007275; P:development; IDA.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; Gonadoliberini.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERIN.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family; Pyrrolidone carboxylic acid.
FT SIGNAL 1 22
FT CHAIN 23 94
FT PEPTIDE 23 32
FT PEPTIDE 36 94
FT MOD_RES 23 23
FT MOD_RES 32 32
FT CONFLICT 86 94
SQ SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;
PROGONADOLIBERIN I.
GONADOLIBERIN I.
GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
PYRROLIDONE CARBOXYLIC ACID.
AMIDATION (G-33 PROVIDE AMIDE GROUP).
ENGRTFKK-->KMDIGHSRNERFL (IN REF. 1).
Query Match 32.1%; Score 52; DB 1; Length 94;
Best Local Similarity 88.9%; Pred. No. 0.84;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 24 HWSYGLSPG 32

Search completed: March 10, 2004, 09:13:57
JOB time : 6.66926 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 08:58:54 ; Search time 30.6381 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGRPGSSGPSQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mmc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	76	45.9	1310 2 Q93N27
2	59	36.4	91 13 Q9PRH0
3	53.5	33.0	374 16 Q92YR6
4	53	32.7	308 16 Q883K3
5	53	32.7	358 17 Q8TU8
6	52.5	32.4	214 16 Q8E7H5
7	52.5	32.4	214 16 Q8E219
8	52.5	32.4	437 15 Q8XKM3
9	52	32.1	64 13 Q8JIF3
10	52	32.1	64 13 Q8JIF2
11	52	32.1	68 13 Q8JIF4
12	52	32.1	87 13 Q9YI26
13	52	32.1	94 13 Q8JFY3
14	52	32.1	96 13 Q8UW80
15	52	32.1	98 13 Q805A5
16	52	32.1	120 13 Q7T059

Q8YV98 anabaena sp
Q8RQ1 streptomyc
Q88902 rattus norv
Q81GA9 arabidopsis
Q88848 eminthopsis
Q80G51 yaba monkey
Q8TSG3 methanosarc
Q9S103 arabidopsis
Q8420 arabidopsis
Q90W09 oncarhynchu
Q8G4A3 bifidobacte
Q855P9 mycobacteri
P72950 synchocyst
Q8CIW4 mus musculu
Q8UQ22 mycobacteri
Q8UQ22 methanosarc
Q8VYH6 arabidopsis
Q8FDB7 escherichia
Q83Q25 shigella fl
Q83B26 coxiella bu
Q50210 melittangui
Q96671 drosophila
Q9VU33 drosophila
Q8TRV2 methanosarc
Q8P8W9 xanthomonas
Q8PKS0 xanthomonas
Q9P6E3 neurospora

ALIGNMENTS

RESULT 1			
ID	Q93N27	PRELIMINARY;	PRT; 1310 AA.
AC	Q93N27;		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Tetanus toxin (Fragment)		
OS	Clostridium tetani.		
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;		
OC	Clostridium		
OX	NCBI_TaxID=1513;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Shumin Z., Dianliang L.;		
RT	"Cloning and sequence analysis of tetanus toxin gene."		
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF389424; AAK72964.2; "		
DR	GO; GO:0004866; Fendopeptidase inhibitor activity; IEA.		
DR	GO; GO:0008237; Fimetalloprotease activity; IEA.		
DR	GO; GO:0015070; F:toxin activity; IEA.		
DR	GO; GO:0008270; F:zinc ion binding; IEA.		
DR	GO; GO:0009405; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR008985; ConA like lec_gl.		
DR	InterPro; IPR001064; Crystallin.		
DR	InterPro; IPR002160; Kunitz_legume.		
DR	InterPro; IPR000395; Peptidase_M27.		
DR	InterPro; IPR006025; Pept_M_Zn_BS.		
DR	Pfam; PF01742; Peptidase M27; 1.		
DR	PRINTS; PR00760; BONTOKILYSIN.		
DR	ProDom; PD001963; Bontoxilysin; 1.		
DR	PROSITE; PS00225; CRYSTALLIN BETAGAMMA; 1.		
DR	PROSITE; PS00442; ZINC_PROTEASE; 1.		
FT	NON_TER 1310 1310		
FT	NON_TER 1310 1310		
SQ	SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;		
Query Match 46.9%; Score 76; DB 2; Length 1310;			

Best Local Similarity 93.8%; Pred. No. 0.04; Mismatches 1; Indels 0; Gaps 0;

QY 16 LQYIKANSKFGITEL 31
DB 830 MOYIKANSKFGITEL 845

RESULT 2

Q9PRHO PRELIMINARY; PRT; 91 AA.
AC Q9PRHO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2003 (TrEMBLrel. 13, Last sequence update)
DE Prepro-mGnRH precursor (Gonadoliberin) (Gonadotropin-releasing hormone) (LH-RH) (Luliberin).
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7937;
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC TISSUE=Brain;
RA "Okubo K., Suetake H., Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GnRH.";
RL Zool. Sci. 16:471-478 (1999).
RN SEQUENCE 374 AA; 40999 MW; 0369AA67790B23D6 CRC64;

Query Match 33.0%; Score 53.5; DB 16; Length 374;

Best Local Similarity 44.8%; Pred. No. 21; Mismatches 13; Conservative 5; Indels 10; Gaps 1;

QY 2 HWSYGLRPGSGPSLQYIKANSKFGITGE 30
DB 21 HWSYGLRPGSGPSLQYIKANSKFGITGE 48

RESULT 4

Q983K3 PRELIMINARY; PRT; 308 AA.
AC Q983K3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycosyl transferase, group 2 family protein.
GN PSPT02354.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN SEQUENCE FROM N.A.
RP STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016864; AA055865.1; -;
DR TIGR; PSPT02354; -;
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF00535; Glycos transf 2; I.
KW Transferase; Complete_proteome.
SQ SEQUENCE 308 AA; 34880 MW; 11732AAE5C2B150 CRC64;

Query Match 32.7%; Score 53; DB 16; Length 308;
Best Local Similarity 36.0%; Pred. No. 20; Mismatches 9; Conservative 5; Indels 11; Gaps 0;

Query Match 36.4%; Score 59; DB 13; Length 91;
Best Local Similarity 66.7%; Pred. No. 0.66; Mismatches 0; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLQ 17
DB 24 HWSYGLRPGSGKRGADSLQ 41

RESULT 3

Q92YR6 PRELIMINARY; PRT; 374 AA.
AC Q92YR6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative muconate cycloisomerase (EC 5.5.1.1).
GN RA0797 OR SNA1461.

QY 3 WYGLRPGSGPSLQVYKANSKFIG 27
DB 234 WYETFGGIGGGMRYVKSELKFLG 258

RESULT 5
Q87ZU8 PRELIMINARY; PRT; 358 AA.
AC Q87ZU8; 32.7%; Score 53; DB 17; Length 358;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Glycosyl transferase.
GN PF1885.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010283; AAL82009.1; -
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR001296; Glyco trans 1.
DR Pfam; PF00534; Glycos transf_1; I.
DR Transferase, Complete proteome.
SQ SEQUENCE 358 AA; 40003 MW; 4D1FAA206B9F8A3C CRC64;

Query Match 32.7%; Score 53; DB 17; Length 358;
Best Local Similarity 52.6%; Pred. No. 24;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 12 SGPSLQVYKANSKFIGITE 30
DB 214 SGEMLPFLKQAKPLGIEE 232

RESULT 6
Q8E7H5 PRELIMINARY; PRT; 214 AA.
AC Q8E7H5; 32.4%; Score 52.5; DB 16; Length 214;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN GS0179.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM315 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Ruznlok C., Buchrieser C., Chevallier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766844; CAD45824.1; -
DR SAGALst; GS0179;
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006402; HAD-SF-IA-V3.
DR InterPro; IPR005833; Hignase/hydrolase.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00702; Hydrolase; 1.

DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-V3; 1.
KW Complete proteome.
SQ SEQUENCE 214 AA; 24470 MW; 7EE27E7BDE946854 CRC64;

Query Match 32.4%; Score 52.5; DB 16; Length 214;
Best Local Similarity 42.4%; Pred. No. 16;
Matches 14; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

QY 2 HW----SYGLRPGSGPSLQVYKANSKFIGITE 30
DB 96 HWLHGHGYRLAVASSPMVD-IKKNLKELGVTE 127

RESULT 8
Q8XKM3 PRELIMINARY; PRT; 437 AA.
AC Q8XKM3; 32.4%; Score 52.5; DB 16; Length 214;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Probable sodium-and chloride-dependent transporter.

DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-V3; 1.
KW Complete proteome.
SQ SEQUENCE 214 AA; 24470 MW; 7EE27E7BDE946854 CRC64;

Query Match 32.4%; Score 52.5; DB 16; Length 214;
Best Local Similarity 42.4%; Pred. No. 16;
Matches 14; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

QY 2 HW----SYGLRPGSGPSLQVYKANSKFIGITE 30
DB 96 HWLHGHGYRLAVASSPMVD-IKKNLKELGVTE 127

RESULT 7
Q8E219 PRELIMINARY; PRT; 214 AA.
AC Q8E219; 32.4%; Score 52.5; DB 16; Length 214;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hydrolase, haloacid dehalogenase-like family.
GN SAG0181.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Masiagnani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014200; AAK99088.1; -
DR TIGR; SAG0181; -
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006402; HAD-SF-IA-V3.
DR InterPro; IPR005833; Hignase/hydrolase.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-V3; 1.
KW Complete proteome.
SQ SEQUENCE 214 AA; 24470 MW; 7EE27E7BDE946854 CRC64;

Query Match 32.4%; Score 52.5; DB 16; Length 214;
Best Local Similarity 42.4%; Pred. No. 16;
Matches 14; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

QY 2 HW----SYGLRPGSGPSLQVYKANSKFIGITE 30
DB 96 HWLHGHGYRLAVASSPMVD-IKKNLKELGVTE 127

Wed Mar 10 10:34:22 2004

GN CP21371.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=13 / Type A;
RC MEDLINE=21664373; PubMed=11728842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayaishi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater";
RJ Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AF003190; BAB107.1; C: integral to plasma membrane; IEA.
DR GO; GO:0005887; C: integral to plasma membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005328; P: neurotransmitter; sodium symporter activity; IEA.
DR GO; GO:0006836; P: neurotransmitter transport; IEA.
DR InterPro; IPR00175; Na/ntran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD004448; Na/ntran_symport; 1.
DR PROSITE; PS0267; NA_NEUROTRAN_SYMPT_3; 1.
KW Complete proteome.
SQ SEQUENCE 437 AA; 47945 MW; AB012AC43DA9A813; CRC64;

Query Match 32.4%; Score 52.5; DB 13; Length 437;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;
QY 3 MSYGLRPGSSPSLYIKANKKFIQI 28
DB 273 PAYGLNP-SSGSLFLITLPKIFQGM 297

RESULT 9
Q8JIF3 PRELIMINARY; PRT; 64 AA.
AC Q8JIF3; 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Gonadotropin-releasing hormone (Fragment).
OS Dentex dentex.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
OC Sparidae; Dentex.
OX NCBI_TaxID=94951;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Elmesiry G.E.;
RT "Sequence comparison of GnRH genes in closely-related Sparidae
RT fishes";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089313; BAC07229.1; --
DR GO; GO:0005576; C: extracellular; IEA.
DR GO; GO:0005183; P: luteinizing hormone-releasing factor activity; IEA.
DR GO; GO:0007275; P: development; IEA.
DR InterPro; IPR002012; GnRH.
DR PRINTS; PR01541; GONADOLIBERIN.
DR PROSITE; PS00473; GnRH; 1.
FT NON_TER 1 64
FT SEQUENCE 64 AA; 7028 MW; 1F91FFB2B4B56D0 CRC64;
Query Match 32.1%; Score 52; DB 13; Length 64;
Best Local Similarity 88.9%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 16 HWSYGLSPG 24
RESULT 10
Q8JIF2 PRELIMINARY; PRT; 64 AA.
AC Q8JIF2; 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Gonadotropin-releasing hormone (Fragment).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Elmesiry G.E.;
RT "Sequence comparison of GnRH genes in closely-related Sparidae
RT fishes";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089314; BAC07230.1; --
DR GO; GO:0005576; C: extracellular; IEA.
DR GO; GO:0005183; P: luteinizing hormone-releasing factor activity; IEA.
DR GO; GO:0007275; P: development; IEA.
DR InterPro; IPR004079; Gonadoliberin.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERIN.
DR PROSITE; PS00473; GnRH; 1.
FT NON_TER 1 64
FT SEQUENCE 64 AA; 6967 MW; 119F0D0CD14BB6D0 CRC64;

Query Match 32.1%; Score 52; DB 13; Length 64;
Best Local Similarity 88.9%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 16 HWSYGLSPG 24
RESULT 11
Q8JIF4 PRELIMINARY; PRT; 68 AA.
AC Q8JIF4; 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Gonadotropin-releasing hormone (Fragment).
OS Acanthopagrus latus (Yellowfin porgy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
OC Sparidae; Acanthopagrus.
OX NCBI_TaxID=8177;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Elmesiry G.E.;
RT "Sequence comparison of GnRH genes in closely-related Sparidae
RT fishes";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089312; BAC07228.1; --
DR GO; GO:0005576; C: extracellular; IEA.
DR GO; GO:0005183; P: luteinizing hormone-releasing factor activity; IEA.
DR GO; GO:0007275; P: development; IEA.
DR InterPro; IPR004079; Gonadoliberin.
DR

DR Pfam: PF00446; GnRH, 1.
 DR PRINTS; PRO1541; GONADOLIBERNI.
 DR PROSITE; PS00473; GNRH; 1.
 FT NON TER 1 1
 FT NON TER 68 68
 FT NON TER 68 68
 SQ SEQUENCE 68 AA; 7543 MW; 067708609FE8E771 CRC64;
 Query Match 32.1%; Score 52; DB 13; Length 68;
 Best Local Similarity 88.9%; Pred. No. 5.2; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 20 HWSYGLSPG 28
 RESULT 12
 QYI26 PRELIMINARY; PRT; 87 AA.
 AC QYI26
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE SGNRH (Gonadoliberein) (Gonadotropin-releasing hormone) (LH-RH)
 DE (Luliberin) (fragment).
 OS Sparus aurata (gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 OX NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC -!- SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 EMBL; AF046801; AAD02427.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadolibereinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PRO1541; GONADOLIBERNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone.
 FT NON TER 1 1
 FT NON TER 87 87
 FT NON TER 87 87
 SQ SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;
 Query Match 32.1%; Score 52; DB 13; Length 87;
 Best Local Similarity 88.9%; Pred. No. 6.8; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 22 HWSYGLSPG 30
 RESULT 13
 QJFY3 PRELIMINARY; PRT; 94 AA.
 AC QJFY3
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE GnRH1 preprohormone precursor (Seabream-type gonadotropin-releasing
 DE hormone precursor).
 GN GNRH1.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Farahmand H., Rahman M.A., Sohn F., Hwang G.-L., Maclean N.;
 RT "Isolation and Expression of Tilapia (Oreochromis niloticus) Serine 8
 RT Type GnRH Coding and Regulatory Sequences";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Sato H., Sakuma Y., Parhar I.S.;
 RT "Molecular cloning of three kinds of GnRH genes and 5' untranslated
 RT regions in tilapia (Oreochromis niloticus)";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF47291; AAM90220.1; -.
 DR EMBL; AB104861; BAC85154.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadolibereinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PRO1541; GONADOLIBERNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 32 GONADOTROPIN-RELEASING HORMONE.
 FT CHAIN 36 94 GNRH-ASSOCIATED PEPTIDE.
 FT CHAIN 94 94
 SQ SEQUENCE 94 AA; 10396 MW; E57DBA832FC078D7 CRC64;
 Query Match 32.1%; Score 52; DB 13; Length 94;
 Best Local Similarity 88.9%; Pred. No. 7.4; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 24 HWSYGLSPG 32
 RESULT 14
 Q8UW80 PRELIMINARY; PRT; 96 AA.
 ID Q8UW80
 AC Q8UW80
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Seabream-type gonadotropin-releasing hormone precursor (Gonadoliberein)
 DE (GnRH) (LH-RH) (Luliberin).
 OS Verasper moseri (barfin flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Pleuronectiformes;
 OC Pleuronectoidae; Pleuronectidae; Verasper.
 OX NCBI_TaxID=98923;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Amato M.;
 RT "Molecular cloning of three cDNAs encoding GnRH in the brain of barfin
 RT flounder";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC -!- SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR EMBL; AB066360; BAB3984.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadolibereinI.
 DR Pfam; PF00446; GNRH; 1.